

Abs, Hermann Josef; Eckert, Thomas; Anderson-Park, Eva  
**Effektivität der Qualifizierung von Teach First Fellows. Abschlussbericht zur  
summativen Evaluation der Sommerakademie von Teach First Deutschland**

*Frankfurt am Main : Gesellschaft zur Förderung Pädagogischer Forschung ; Deutsches Institut für  
Internationale Pädagogische Forschung 2016, 106 S. - (Materialien zur Bildungsforschung; 34)*



Quellenangabe/ Reference:

Abs, Hermann Josef; Eckert, Thomas; Anderson-Park, Eva: Effektivität der Qualifizierung von Teach First Fellows. Abschlussbericht zur summativen Evaluation der Sommerakademie von Teach First Deutschland. Frankfurt am Main : Gesellschaft zur Förderung Pädagogischer Forschung ; Deutsches Institut für Internationale Pädagogische Forschung 2016, 106 S. - (Materialien zur Bildungsforschung; 34) - URN: urn:nbn:de:0111-pedocs-124697 - DOI: 10.25656/01:12469

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-124697>

<https://doi.org/10.25656/01:12469>

in Kooperation mit / in cooperation with:



**GFPF**

Gesellschaft zur Förderung  
Pädagogischer Forschung e.V.

<http://www.gfpf.info>

**Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

**Kontakt / Contact:**

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Mitglied der

  
Leibniz-Gemeinschaft



Hermann Josef Abs • Thomas Eckert • Eva Anderson-Park

## Effektivität der Qualifizierung von Teach First Fellows

Abschlussbericht zur summativen Evaluation  
der Sommerakademie von Teach First Deutschland

Hermann Josef Abs / Thomas Eckert / Eva Anderson-Park

## Effektivität der Qualifizierung von Teach First Fellows

Abschlussbericht zur summativen Evaluation  
der Sommerakademie von Teach First Deutschland

Materialien zur Bildungsforschung

Frankfurt am Main 2016

## **Fachbeirat der Materialien zur Bildungsforschung**

MR'in Cäcilie Daumen, Ministerium für Bildung, Frauen und Jugend, Mainz

Prof. em. Dr. Jörg Schlömerkemper, Göttingen

Dr. Marius Gerecht, Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt am Main

Prof. Dr. Eckhard Klieme, Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt am Main

Dr. Renate Martini, Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt am Main

Dr. Simon Rettelbach, Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt am Main



### **Adressen der Autoren:**

Prof. Dr. Hermann Josef Abs  
Universität Duisburg Essen  
Fakultät Bildungswissenschaften  
Universitätsstr. 2  
45141 Essen

Prof. Dr. Thomas Eckert  
Ludwig-Maximilians-Universität München  
Fakultät für Psychologie und Pädagogik  
Leopoldstr. 13  
80802 München

Eva Anderson-Park  
Universität Duisburg Essen  
Fakultät Bildungswissenschaften  
Universitätsstr. 2  
45141 Essen

Unter Mitarbeit von: Eva Blomberg, Ludwig-Maximilians-Universität München, Fakultät für Psychologie und Pädagogik, Leopoldstr. 13, 80802 München

### **Bibliografische Information der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der deutschen Nationalbibliographie; detaillierte bibliografische Daten sind im Internet über <http://d-nb.de> abrufbar.

Copyright 2016 by

Gesellschaft zur Förderung Pädagogischer Forschung;  
Deutsches Institut für Internationale Pädagogische Forschung,  
Schloßstraße 29, D-60486 Frankfurt am Main.

Printed in Germany

ISBN: 978-3-923638-55-0

Materialien zur Bildungsforschung, Bd. 34



## Inhaltsverzeichnis

1	Executive Summary .....	7
2	Ziele der Evaluationsstudie .....	10
3	Erhebungsdesign .....	12
4	Entwicklung der Erhebungsinstrumente.....	13
5	Durchführung der Untersuchung .....	15
5.1	Kontinuität in der Zusammensetzung der Stichprobe .....	15
6	Persönliche Voraussetzungen der Fellows.....	18
7	Prozesse des Kompetenzerwerbs .....	22
7.1	Merkmale der Lernumgebung.....	22
	Zweite Erhebung: Einschätzung der Qualität des Online Campus.....	22
7.1.1	Gestaltungsqualität der Lernplattform .....	22
7.1.2	Interaktion zwischen den Fellows und den Trainern / Trainerinnen.....	23
7.1.3	Qualität der Lehrmaterialien.....	24
	Dritte Erhebung: Einschätzung der Praxisteile der Sommerakademie.....	27
7.1.4	Seminarteil Katlenburg.....	28
7.1.5	Känguru Erfahrung .....	29
7.1.6	MSact! .....	31
7.1.7	Offene Kommentare der Befragten zur Sommerakademie.....	32
	Vierte Erhebung: Einschätzung der schuleinsatzbegleitenden Maßnahmen .....	35
7.1.8	Balkontreffen.....	35
7.1.9	Trainer-Fellow-Treffen .....	36
7.1.10	Fortbildungen .....	37
7.1.11	Hospitation/Bilanzgespräch .....	38
7.1.12	Kompetenzmatrix.....	39
7.1.13	TFD Werkstatt .....	40
7.1.14	Lernumgebung Schule.....	42
7.1.15	Offene Kommentare zur Qualifizierung während des Schuleinsatzes.....	45
	Trainingsübergreifende Merkmale der Lernumgebung.....	49
7.1.16	Einschätzung der Qualität der Trainer und Trainerinnen .....	49
7.1.17	Kohärenz in der Ausbildung .....	51

7.2	Merkmale der Nutzung der Lernumgebung.....	52
7.2.1	Nutzung von Lernstrategien.....	52
7.2.2	Professionelle Kooperation .....	53
7.2.3	Zeitmanagement .....	54
7.2.4	Nutzung der TFD Werkstatt .....	55
8	Ergebnisse .....	56
8.1	Selbstwirksamkeitserwartung .....	56
8.2	Selbsteingeschätzte Unterrichtskompetenz .....	60
8.3	Testergebnisse.....	64
8.4	Professionalität von TFD Fellows im Vergleich.....	74
8.4.1	TFD Fellows im Vergleich: Die TEDS-M Studie .....	74
8.4.2	TFD-Fellows im Vergleich mit Stipendiaten der Stiftung der Deutschen Wirtschaft.....	77
9	Fazit .....	80
10	Ausblick .....	82
11	Literaturverzeichnis.....	84
12	Anhang .....	86

## Abbildungsverzeichnis

Abbildung 1: Hypothetisches Wirkmodell .....	10
Abbildung 2: Erhebungsplan für die Evaluation des TFD Trainingsprogramms .....	12
Abbildung 3: Zusammensetzung der Stichprobe über die vier Erhebungswellen.....	16
Abbildung 4: Selbstwirksamkeitserwartungen über drei Messzeitpunkte.....	60
Abbildung 5: Durchschnittliche selbsteingeschätzte Unterrichtskompetenz über drei Messzeitpunkte .....	63
Abbildung 6: Beispieltestitem zu Lerninhalten aus Woche 2 des Online Campus .....	65
Abbildung 7: Mittlere Lösungshäufigkeit der Testfragen für TFD Fellows über drei Erhebungszeitpunkte.....	72
Abbildung 8: Erreichte Testpunkte im Wissenstest über drei Messzeitpunkte .....	73
Abbildung 9: Wissensfragen aus PUW-Test der TEDS-M Studie - Lösungshäufigkeit über drei Messzeitpunkte.....	88
Abbildung 10: Wissensfragen zu TFD-Inhalten: Lösungshäufigkeit über 3 Messzeitpunkte...	88

## Tabellenverzeichnis

Tabelle 1:	Übersicht der erhobenen Konstrukte nach Messzeitpunkten .....	14
Tabelle 2:	Rollenselbstbild der Fellows - Deskriptive Statistiken .....	20
Tabelle 3:	Belastungserleben - Deskriptive Statistiken .....	21
Tabelle 4:	Gestaltungsqualität Lernplattform - Deskriptive Statistiken der Einzelitems ....	23
Tabelle 5:	Interaktion zwischen den Beteiligten – Deskriptive Statistiken .....	24
Tabelle 6:	Qualität der Lehrmaterialien: Menge des Inputs, Strukturiertheit, Verständlichkeit – Deskriptive Statistiken .....	25
Tabelle 7:	Qualität der Lehrmaterialien: Praxisorientierung – Deskriptive Statistiken .....	26
Tabelle 8:	Qualität der Lehrmaterialien: Kognitive Aktivierung – Deskriptive Statistiken..	26
Tabelle 9:	Qualität der Lehrmaterialien: Wissenschaftliche Fundierung – Deskriptive Statistiken.....	27
Tabelle 10:	Professionelle Kooperation der Fellows während des Seminarteils auf der Katlenburg - Deskriptive Statistiken .....	28
Tabelle 11:	Gestaltung des Trainingsteils "Seminarteil auf der Katlenburg" - Deskriptive Statistiken .....	29
Tabelle 12:	Erfahrungen während des Trainingsteils "Känguru" - Deskriptive Statistiken ...	30
Tabelle 13:	Erfahrungen mit dem Trainingsteil MSact! - Deskriptive Statistiken .....	31
Tabelle 14:	Einschätzung der TFD-Balkon-Treffen - Deskriptive Statistiken .....	36
Tabelle 15:	Einschätzung der TFD-Trainer-Fellow-Treffen - Deskriptive Statistiken.....	37
Tabelle 16:	Einschätzung der TFD-Fortbildungen - Deskriptive Statistiken .....	38
Tabelle 17:	Einschätzung der TFD-Kompetenzmatrix - Deskriptive Statistiken .....	39
Tabelle 18:	Nutzerfreundlichkeit der TFD-Werkstatt .....	40
Tabelle 19:	Einschätzung der sozialen Eingebundenheit mit versus ohne Betreuungsperson - Deskriptive Statistiken .....	43
Tabelle 20:	Lerngelegenheiten/Einsatzprofil der Fellows - Deskriptive Statistiken.....	45
Tabelle 21:	Einschätzung der Trainer und Trainerinnen - Deskriptive Statistiken .....	50
Tabelle 22:	Einschätzung der Kohärenz der Ausbildung – Deskriptive Statistiken .....	52
Tabelle 23:	Nutzung von Lernstrategien – Deskriptive Statistiken .....	53
Tabelle 24:	Professionelle Kooperation – Deskriptive Statistiken.....	53
Tabelle 25:	Zeitmanagement Online Campus - Deskriptive Statistiken .....	54
Tabelle 26:	Zeitmanagement Seminarteil/ MSact!/ Schuleinsatz - Deskriptive Statistiken..	54
Tabelle 27:	Nutzung der TFD Werkstatt - Deskriptive Statistiken .....	55
Tabelle 28:	Selbstwirksamkeitserwartung - Klassenführung - Deskriptive Statistiken .....	57
Tabelle 29:	Selbstwirksamkeitserwartung – Unterrichtsstrategien - Deskriptive Statistiken .....	58
Tabelle 30:	Selbstwirksamkeitserwartung - Schülermotivation - Deskriptive Statistiken ....	59
Tabelle 31:	Selbsteingeschätzte Unterrichtskompetenz über drei Messzeitpunkte - Deskriptive Statistiken .....	61
Tabelle 32:	Kompetenzeinschätzung Einzelitems "Schuleinsatz" - Deskriptive Statistiken ..	63
Tabelle 33:	Testergebnisse im Bereich Unterrichten – Planen .....	66

Tabelle 34:	Testergebnisse im Bereich Unterrichten – Umsetzen .....	67
Tabelle 35:	Testergebnisse im Bereich Leadership .....	68
Tabelle 36:	Testergebnisse im Bereich Leadership - Zusatzfragen .....	69
Tabelle 37:	Testergebnisse im Bereich Diagnose .....	70
Tabelle 38:	Testergebnisse im Bereich Sprachförderung.....	71
Tabelle 39:	Testergebnisse im Bereich TFD-Einsatz .....	72
Tabelle 40:	Testpunkte - Deskriptive Statistiken .....	74
Tabelle 41:	Vergleich der Testergebnisse zum Pädagogischen Unterrichtswissen von Lehramtsstudierenden und TFD Fellows – Lernausgangslage.....	75
Tabelle 42:	Vergleich der Testergebnisse zum Pädagogischen Unterrichtswissen von Referendaren und TFD Fellows - Leistung am Ende der Ausbildung.....	77
Tabelle 43:	Vergleich der Testergebnisse zum Pädagogischen Unterrichtswissen von TFD Fellows und SDW Stipendiaten .....	78
Tabelle 44:	Vergleich der Testergebnisse der TFD Wissensfragen von TFD Fellows und SDW Stipendiaten .....	79
Tabelle 45:	Deskriptive Statistiken .....	86
Tabelle 46:	Kategorisierung der Studienabschlüsse der TFD Fellows .....	87
Tabelle 47:	I D-Nummern der Personen, die nicht in allen Befragungswellen enthalten sind .....	88
Tabelle 48:	Ausführliche Testergebnisse im Bereich Unterrichten – Planen .....	90
Tabelle 49:	Ausführliche Testergebnisse für den Bereich Unterrichten – Umsetzen .....	93
Tabelle 50:	Ausführliche Ergebnisse im Bereich Leadership .....	96
Tabelle 51:	Ausführliche Ergebnisse für den Bereich Diagnose .....	98
Tabelle 52:	Ausführliche Ergebnisse im Bereich Sprachförderung .....	100



## 1 Executive Summary

### *Kontext und Ziel der Studie*

*Teach First Fellows* nehmen wie Lehrkräfte eine hohe Verantwortung für die weitere Entwicklung von Kindern und Jugendlichen wahr. Anders als Lehrkräfte verfügen sie jedoch nicht über eine grundständige Ausbildung für ihre spezifische Tätigkeit. Deshalb stellt sich die Frage, inwiefern es gelingen kann, durch eine sorgfältige Auswahl der *Fellows*, sowie durch eine intensive Kurzqualifizierung gute Voraussetzungen für den Erfolg ihrer Tätigkeit zu schaffen.

Die vorliegende Studie evaluiert die Ergebnisse des Qualifizierungsprogramms, das *Teach First Deutschland* für seine *Fellows* entwickelt und implementiert hat. Dabei interessiert insbesondere, welches Wissen zur Unterstützung von Lernprozessen, welche Fähigkeiten zur Gestaltung von Unterricht und welche Erwartung von pädagogischer Wirksamkeit innerhalb eines halben Jahres bei den *Fellows* aufgebaut werden können. Um ein tieferes Verständnis der Wirkweise des Qualifizierungsprogramms zu ermöglichen, wurden neben den Ergebnissen auch dessen Prozesskomponenten einer Bewertung unterzogen.

### *Planung und Verlauf der Untersuchung*

Die Untersuchung basiert auf einem Wirkungsmodell, das Bedingungen, Prozesse und intendierte Ergebnisse der Qualifizierung auf der institutionellen und individuellen Ebene ausdifferenziert. Ausgehend von dieser Konzipierung wurden Untersuchungsinstrumente teilweise neu entwickelt, teilweise aus anderen Untersuchungen adaptiert und teilweise identisch übernommen. Letzteres ermöglicht in der Analyse der Ergebnisse den Vergleich mit Ergebnissen anderer Untersuchungen.

Die Studie setzt ein Längsschnittdesign um, d.h. die *Fellows* wurden im Verlauf von acht Monaten vier Mal befragt. Die erste Befragung vor dem Start der *Blended Learning* Phase, dem *Online Campus* erfasst die Vorkenntnisse und Vorerfahrungen der *Fellows*, die zweite Befragung zum Abschluss des *Online Campus* dient dazu die *Blended Learning* Phase zu bewerten und den erreichten Wissenstand zu dokumentieren. In zwei weiteren Befragungen, nach den Präsenz- und Praxisphasen vor und ca. vier Monate nach dem Beginn des Schuleinsatzes werden die Stabilität und weitere Entwicklung von Wissen und pädagogischer Handlungsfähigkeit in den Blick genommen.

### *Merkmale der Fellows im Qualifizierungsprogramm*

Insgesamt nahmen 65 *Fellows* an der Studie teil. Von 46 Teilnehmenden liegen Daten aus der ersten und letzten Welle vor. Die Differenz begründet sich nicht nur aus den üblichen Schwundquoten in Panelstudien, sondern auch daraus, dass *Fellows* auch noch nach dem ersten Erhebungszeitpunkt für das Programm zugelassen wurden (Nachrekrutierung). Die *Fellows* rekrutierten sich aus einer großen Breite von Studiengängen, sie sind zu 60% weiblich und im Mittel 27,8 Jahre alt. Bei etwa einem Fünftel der *Fellows* verweist die Sprache in der Herkunftsfamilie auf einen Migrationshintergrund. Dies und die vorhandenen Auslandsaufenthalte von mindestens einem Jahr bei mehr als der Hälfte der *Fellows* deuten

auf Erfahrungen im Umgang mit unterschiedlichen kulturellen Horizonten. Ebenso können über vier Fünftel der *Fellows* auf Erfahrungen in der Arbeit mit Kindern und Jugendlichen vor der Qualifizierung zurückgreifen. Weiterhin weisen die *Fellows* entsprechend dem rigorosen Auswahlverfahren überdurchschnittlich gute Noten sowohl im Abitur als auch in ihren Studienabschlüssen auf; dies lässt auf eine hohe allgemeine Leistungsfähigkeit schließen.

### *Bewertung von konstitutiven Elementen des Qualifizierungsprogramms*

- Die eingesetzte *Blended Learning* Umgebung wurde hinsichtlich dreier Aspekte untersucht: Technische Umsetzung, Interaktion zwischen Teilnehmenden und mit Trainer/inne/n, sowie inhaltliche Qualität der Lernmaterialien.
- Die Bewertung der Kontakt- und Praxisphasen vor dem Schuleinsatz umfasst die folgenden drei Komponenten: Die Präsenzphasen der Seminarwochen in einem Bildungshaus (Katlenburg), die Praktika bei *Fellows* der beiden vorherigen Jahrgänge in deren Einsatzschulen, sowie Ferienschulen für bildungsbenachteiligte Kinder.
- Die Bewertung der begleitenden Qualifizierung während des Schuleinsatzes bezieht sich auf ein Netz von Maßnahmen, die kriteriengeleitete Reflexion und fortschreitende Kompetenzentwicklung absichern sollen. Dazu zählten regelmäßige Trainerbesuche, bei denen die pädagogische Praxis der *Fellows* anhand einer Kompetenzmatrix bewertet wird, individuelle Gespräche zwischen *Fellows* und Trainern, themenbezogene Fortbildungen in Gruppen, sowie Reflexionstreffen auf der Ebene der *Fellows* (*peer level*) stattfinden. Zusätzlich erfasste die Studie auch die Einsatzschulen als Raum der Kooperation und professionsbezogenen Sozialisation.

Insgesamt weisen sämtliche Elemente der Qualifizierung eine hohe Akzeptanz auf. Positive Bewertungen in Fragebogenskalen zeigen sich sowohl für die gewählten Formate (Online-Campus, Präsenzphasen, Praxiserprobungen, Austausch zwischen *Fellows*, Mentoring und Weiterbildung während des Einsatzes) als auch für die Inhalte, Kohärenz der Materialien, und Interaktion mit den Trainern. Einzelne Hinweise auf Entwicklungspotential erscheinen in offenen Antwortformaten und betreffen konkrete Teilaspekte der technischen Umsetzung, einzelne Inhalte und den Wunsch nach mehr Praxisanleitungen in spezifischen Problemstellungen.

### *Effekte der Qualifizierung*

Aus einer psychometrischen Perspektive stellt das getestete Wissen den härtesten Indikator für Erfolg dar. Der Testwert für das Wissen ist weniger als selbst eingeschätzte Kompetenz und Wirkungseinschätzungen durch mögliche Beurteilungsfehler verzerrt. Gleichwohl sind die Selbsteinschätzungen wichtige Ergänzungen zum Test, weil sie die subjektive Perspektive einfangen und über die kognitiven Effekte bei den *Fellows* hinausweisen auf deren künftiges Verhalten und Selbstbewusstsein.

Die kognitiven Ergebnisse der Qualifizierung wurden in zwei Testteilen erfasst. Neben einem Testteil, der spezifisch auf der Grundlage von Materialien aus dem *Online Campus* entwickelt wurde, nutzte ein weiterer Testteil Fragen einer internationalen Studie zu pädagogischem

Unterrichtswissen Dadurch wurde der Vergleich zu einer repräsentativen Stichprobe in Deutschland möglich.

Die Qualifizierung erweist sich als eindeutig funktional im Hinblick auf den intendierten Wissensaufbau. Die Kenntnisse der *Fellows* zu den Inhalten der Qualifizierung haben sich schon im Verlauf der *Online*-Beschäftigung mit den bereitgestellten Materialien verdoppelt. Der Wissensaufbau ist auch mehrere Monate nach der ursprünglichen Aneignungsphase noch stabil. Dieses Ergebnis ist umso bemerkenswerter, als ein paralleler Test mit Fragen aus einer großen Repräsentativbefragung von Lehramtskandidat/inn/en zeigt, dass die *Teach First-Fellows* schon zu Beginn ihrer Qualifizierung ein überdurchschnittliches Wissen zu Schule und Unterricht mitbringen, und ihr pädagogisches Unterrichtswissen im gestellten Testausschnitt nach einem halben Jahr dem von Lehrkräften im Vorbereitungsdienst entspricht.

Ein leichter Rückgang der erwarteten Selbstwirksamkeit bei den *Fellows* kann beobachtet werden, wenn die Werte vor der Qualifizierung und nach den ersten Monaten im Einsatz verglichen werden. Auch wenn sich die *Fellows* zu allen Messzeitpunkten und in allen erfragten Tätigkeitsbereichen als eher wirksam denn unwirksam empfinden, so ist doch insbesondere in Bezug auf Schülermotivation und Klassenführung der Ausblick auf die eigene Wirksamkeit nach einem halben Jahr etwas reduziert. Dieser Befund zeigt einen scheinbaren Kontrast zum gelingenden Wissensaufbau und steht vermutlich mit der vertieften Erfahrung und Kenntnis von real vorfindlichen Bedingungen schulischer Arbeit in Zusammenhang. Diese Entwicklung der Selbstwirksamkeit ist für den ersten längeren Praxiskontakt mit schulischen Handlungsfeldern nicht ungewöhnlich und kann als solche auch in das Qualifizierungsprogramm von *Teach First* einbezogen werden.

Positiv entwickelt sich hingegen im Verlauf der Qualifizierung die wahrgenommene Unterrichtskompetenz der *Fellows*. D.h. die *Fellows* sehen sich im Verlauf mehr und mehr in der Lage alltägliche Anforderungen an die Unterrichtstätigkeit gut zu erfüllen. Eventuell gibt es in diesem Bereich aber noch ein nicht ausgeschöpftes Entwicklungspotential. Die Formulierung „eventuell“ ist dabei bewusst vorsichtig gewählt, weil zwar Vergleichswerte aus größeren Studien existieren, die die Annahme eines solchen Entwicklungspotentials nahe legen, hier jedoch berücksichtigt werden muss, dass die *Fellows* an Schulen in schwieriger Lage arbeiten, in denen beginnende Lehrkräfte vor dem Hintergrund der realen Probleme vielleicht zu einer geringeren selbsteingeschätzten Kompetenz neigen als im Vergleich zur Gesamtheit aller Schulen.



## 2 Ziele der Evaluationsstudie

Die vorliegende Studie bezieht sich auf die Evaluation unterschiedlicher Merkmale des Teach First Deutschland (TFD) Qualifizierungsprogramms und dessen differentielle Effekte auf den Wissens- und Kompetenzerwerb der zu Qualifizierenden (Fellows).

Die Evaluationsstudie ist von einer summativen Zielsetzung geprägt. Das heißt, es ist kein Ziel die Qualifizierungsmaßnahme in ihrem Verlauf zu verändern, sondern es gilt, Effekte der unterschiedlichen Maßnahmen bei den Fellows aufzuzeigen. Dabei geht es sowohl darum, Veränderungen bei den Fellows im Vergleich zu ihrem Vorwissen im Längsschnitt zu dokumentieren als auch darum, das Ausmaß des Kompetenzzuwachses im Vergleich zu anderen Personen zu beschreiben. Der Kompetenzzuwachs wird hier aus Veränderungen von Wissen in einem Wissenstest sowie aus der Veränderung von Kompetenzeinschätzungen und Selbstwirksamkeitserwartungen der Fellows erschlossen.

Die angestrebten Effekte des Trainingsprogramms sind in einer Kompetenzmatrix niedergelegt, die sich vor allem auf drei Bereiche bezieht: Kultur (v.a. Lern- und Feedback-Kultur), Leadership (v.a. Reflexive Selbststeuerung und Verantwortungsübernahme) und Unterricht (Ziele setzen, Diagnose, Qualitätsentwicklung, Praxisreflexion). Diese Kompetenzmatrix bildet den Ausgangspunkt der summativen Evaluation.

Die Evaluationsstudie bezieht sich zur Erklärung von Unterschieden in den Ergebnissen vor allem auf die Prozesse des Kompetenzerwerbs, welche sich in einem hypothetischen Wirkmodell verorten lassen (siehe Abbildung 1).

Voraussetzungen	Prozesse des Kompetenzerwerbs	Ergebnisse
	<ul style="list-style-type: none"> <li>Qualifizierungsweg (Referendariat bzw. TFD-Trainingsprogramm)</li> </ul>	
	Merkmale der Lernumgebung: <ul style="list-style-type: none"> <li>Einschätzung der Qualität der Trainingsteile</li> <li>Einschätzung der Qualität der Trainer/Ausbilder</li> <li>Kohärenz des Trainings</li> <li>Einschätzung der Erfahrung in Schule und Unterricht.</li> </ul>	
<ul style="list-style-type: none"> <li>Geschlecht, Alter</li> <li>Art des Studienabschlusses, Noten</li> <li>Pädagogische Vorerfahrung</li> <li>Auslandsaufenthalt</li> <li>Eingangswissen</li> <li>Anfängliches Selbstkonzept</li> <li>Belastungserleben</li> </ul>	Merkmale der Nutzung der Lernumgebung: <ul style="list-style-type: none"> <li>Nutzung von Lernstrategien</li> <li>Nutzung von Vorbereitungszeit</li> <li>Professionelle Kooperation</li> </ul>	<ul style="list-style-type: none"> <li>Wissenszuwachs</li> <li>Kompetenzzuwachs</li> <li>Selbstwirksamkeitserwartung</li> </ul>

Abbildung 1: Hypothetisches Wirkmodell

Dieses ist in erster Linie ein Input-Prozess-Output Modell, das davon ausgeht, dass Bildungsergebnisse auf Bildungsvoraussetzungen und -prozessen aufbauen. Zusätzlich sind hier die von Fend (2008) beschriebenen Ebenen des Bildungssystems mitberücksichtigt. Fend unterscheidet drei Systemebenen: Die Makroebene der politischen Entscheidungen und Bildungssteuerung, die Mesoebene der Lehrpläne und die Mikroebene der Individuen und der einzelnen Lehr-Lernsituationen. Durch diese Unterscheidung lassen sich Prozesse des Kompetenzerwerbs in Merkmale der Lernumgebung auf der Mesoebene (z.B. Kohärenz des Trainings) und Merkmale der Nutzung der Lernumgebung (z.B. Professionelle Kooperation) auf der Mikroebene trennen.

Um die erreichten Ergebnisse beurteilen zu können, werden die Ergebnisse zum einen mit einer unabhängigen Stichprobe, zum anderen mit Ergebnissen anderer Studien verglichen. Um dies tun zu können, mussten für die Evaluationsstudie z.T. identische Testinstrumente, wie in diesen Vergleichsstudien eingesetzt werden. Die mittelfristige Stabilität der Ergebnisse wurde durch eine Abschlussbefragung fünf Monate nach Abschluss der Sommerakademie ermittelt. Um auch längerfristige Ergebnisse (Outcome) zu untersuchen, könnte eine weitere Erhebung mit der gleichen Stichprobe zu einem späteren Zeitraum (z.B. nach einem weiteren Jahr) vorgenommen werden.

### 3 Erhebungsdesign

Das Erhebungsdesign ist ein Längsschnittdesign mit insgesamt vier Erhebungszeitpunkten. Abbildung 2 visualisiert das gesamte Erhebungsdesign und zeigt wie sich die Befragungen in den Ablauf der TFD Sommerakademie einbetteten.

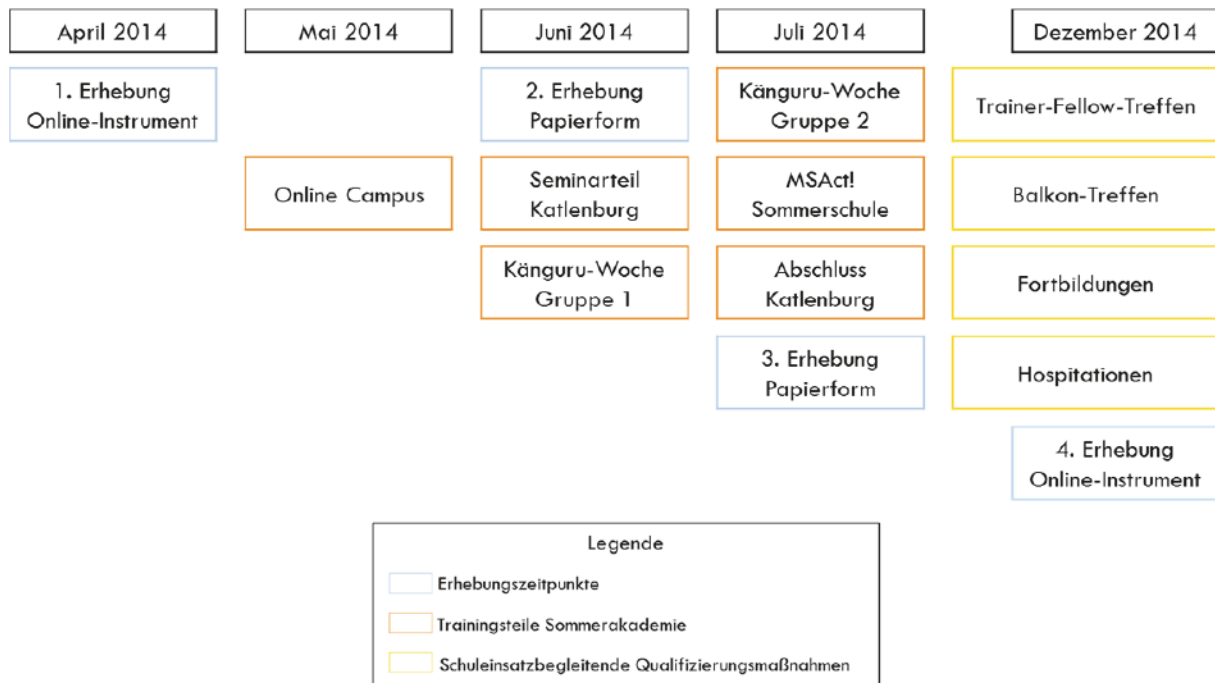


Abbildung 2: Erhebungsplan für die Evaluation des TFD Trainingsprogramms

Der erste Erhebungszeitraum lag vor Beginn (bzw. zur Eröffnung) der Online-Lernphase. Am 28.04.2014 wurde der Zugang zum Online-Fragebogen frei geschaltet und blieb bis zur Durchführung der 2. Welle geöffnet. Der letzte Teilnehmer/die letzte Teilnehmerin hat die Umfrage bereits am 31. Mai ausgefüllt. Nach dem Start der Umfrage wurden in verschiedenen Wellen noch weitere Fellows in die Umfrage aufgenommen. Gleichzeitig wurden Fellows wieder aus dem Daten-Pool herausgenommen, weil sie die Teilnahme am Programm absagten. Die später hinzugefügten, nachrekrutierten Fellows bearbeiteten die Umfrage jedoch alle im selben Zeitraum (also bis zum 31. Mai) wie die zu Beginn angeschriebenen Fellows. Die zweite Erhebung fand am 11.06.2014 im Anschluss an die Online-Lernphase noch vor Beginn des Seminarteils auf der Katlenburg in den Räumen der Burg als Paper-Pencil Befragung statt. Die dritte Erhebung – ebenfalls Paper-Pencil – wurde am 30.07.2014 nach dem Seminarteil auf der Katlenburg und der Sommerschule MSact! erneut in den Räumen der Burg durchgeführt. Die vierte Erhebung erfolgte wieder elektronisch. Der web-basierte Zugang zum Online-Fragebogen bestand vom 08.12.2014 bis zum 12.03.2015. Das ursprünglich kürzer geplante Zeitfenster für diese Erhebung wurde zweimal verlängert, parallel wurden die Fellows, die sich noch nicht beteiligt hatten, zweimal per E-Mail an die anonyme Teilnahmemöglichkeit erinnert.

## 4 Entwicklung der Erhebungsinstrumente

Der Instrumentenentwicklung gingen eine Literaturrecherche zur Wirksamkeit von Lehrer-einführungsprogrammen (Abs, Anderson-Park, 2014), Lehrerkompetenzforschung (Frey, 2014) und Wissenstests in der Lehrerbildungsforschung (Kraus et al., 2008; König et al., 2015; König, 2014, Neuweg, 2014) sowie eine sorgfältige Durchsicht der internen Lehrmaterialien von TFD voraus. Dem hypothetischen Wirkungsmodell folgend sollten sowohl Input-, Prozess-, als auch Output-Variablen erhoben werden. Daher wurden zum einen persönliche Voraussetzungen der Fellows, Einstellungen (z.B. Einstellungen zur Lernkultur) und Werte (z.B. TFD-Werte) erhoben. Zum anderen sollte die Qualität und Kohärenz der einzelnen Trainingsteile erfasst werden. In Einklang mit den Lernzielen des TFD Trainingsprogramms sollten außerdem Output-Variablen wie Wissen und selbsteingeschätzte Kompetenzen im Bereich der Unterrichtsplanung und -gestaltung sowie der Leistungsdiagnose erhoben werden.

Die einzelnen Instrumente zur Erfassung der gewünschten Konstrukte wurden teilweise selbst konzipiert, teilweise wurde auf bereits bestehende Skalen aus der Forschung zurückgegriffen, die dann ggf. adaptiert wurden. Eine Hauptquelle war dabei eine vom Deutschen Institut für internationale pädagogische Forschung (DIPF) im Auftrag des hessischen Kultusministeriums durchgeführte Evaluationsstudie aller hessischen Studienseminare mit dem Titel „Pädagogische Entwicklungsbilanzen mit hessischen Studienseminaren“ (PEB-Sem, vgl. Abs et al., 2009). Dort wurden unter anderem Lehrende im Vorbereitungsdienst zu ihrem Ausbildungsprogramm befragt, sodass wir hier auf eine Anzahl von bereits validierten Skalen zurückgreifen konnten. Des Weiteren wurden die Skalen zur Selbstwirksamkeit (self-efficacy) von Tschannen-Moran und Woolfolk-Hoy (2001) übernommen. Da diese aber nur in der Englischen Originalversion vorlagen, wurde eine wortgetreue Übersetzung ins Deutsche vorgenommen.

Bei der Konzeption des Wissenstests wurden einerseits Items nach Durchsicht der TFD Lehrmaterialien selbst entwickelt. Andererseits wurden, zwecks eines angestrebten Vergleichs mit Ergebnissen aus anderen Studien, inhaltlich nahe liegende Items aus dem Wissenstest „Pädagogisches Unterrichtswissen“ aus der TEDS-M Studie von König und Blömeke (2010) übernommen.

Eine vollständige Übersicht aller erhobenen Konstrukte der vier Erhebungen findet sich in Tabelle 1.



	MZP 1	MZP 2	MZP 3	MZP 4
<b>Erhobene Konstrukte:</b>				
<b>Persönliche Voraussetzungen</b>				
Geschlecht, Alter, Abiturnote, Studium, Abschlussnote, Auslandsaufenthalt, Erfahrung in der Arbeit mit Kindern, Muttersprache, Vorbereitungsdienst, Lehramt	x			
Beginn Schuleinsatz				x
Rollenselbstbild der Fellows		x		
Belastungserleben (PEB-Sem)		x	x	x
<b>Merkmale der Lernumgebung</b>				
Einschätzung der Qualität des Online Campus		x		
Einschätzung Seminarteil Katlenburg			x	
Einschätzung Känguru-Erfahrung			x	
Einschätzung MSact!			x	
Anmerkungen zum Trainingsprogramm			x	
Einschätzung Trainer-Fellow-Treffen				x
Einschätzung TFD-Kompetenzmatrix				x
Einschätzung der TFD- Balkentreffen				x
Einschätzung der TFD-Fortbildungen				x
Einschätzung TFD-Werkstatt				x
Verbesserungsvorschläge zur TFD-Werkstatt				x
Bewertung der Trainer (PEB-Sem adaptiert)			x	x
Kohärenz von Ausbildungsteilen (PEB-Sem)			x	
Zufriedenheit mit Unterstützung durch TFD während Schuleinsatz				x
Verbesserungsvorschläge zum TFD-Qualifizierungsprogramm während Schuleinsatz				x
Lerngelegenheiten (Einsatzprofil an Schule)				x
Qualität der sozialen Eingebundenheit an der Einsatzschule				x
<b>Merkmale der Nutzung der Lernumgebung</b>				
Vorbereitungszeit (PEB-Sem)		x	x	x
Nutzung von Lernstrategien (PEB-Sem)	x			
Soziale Ressourcenorientierung (PEB-Sem)		x		
Nutzung der TFD Werkstatt				x
<b>Ergebnisse</b>				
Testfragen (König und Blömeke, 2010)	x	x	x	x
Selbstwirksamkeit – Unterrichtsstrategien (nach Tschannen-Moran und Woolfolk-Hoy, 2001)	x		x	x
Selbstwirksamkeit – Klassenführung (nach Tschannen-Moran und Woolfolk-Hoy, 2001)	x		x	x
Selbstwirksamkeit – Schülermotivation (nach Tschannen-Moran und Woolfolk-Hoy, 2001)	x		x	x
Kompetenzeinschätzung – Unterrichtsbezogene Tätigkeiten (PEB-Sem)	x		x	x
Kompetenzeinschätzung – Einzelitems (PEB-Sem)	x		x	x
<b>Anzahl der Items insgesamt (excl. Testitems)</b>	<b>73</b>	<b>66</b>	<b>104</b>	<b>137</b>
<b>Anzahl der Testfragen insgesamt:</b>	<b>23 (52)<sup>1</sup></b>	<b>23 (52)</b>	<b>5 (8)</b>	<b>25 (57)</b>
Anzahl Testfragen TFD	15 (28)	15 (28)	5 (8)	17 (33)
Anzahl Testfragen TEDS-M	8 (24)	8 (24)	0	8 (24)

Tabelle 1: Übersicht der erhobenen Konstrukte nach Messzeitpunkten

<sup>1</sup> In Klammern ist die Anzahl an Testfragen unter Berücksichtigung der Teilfragen angegeben.

## 5 Durchführung der Untersuchung

Die Online-Instrumente der ersten und vierten Welle wurden mit dem Programm IBM SPSS Data Collection erstellt; über dieses Programm wurde anschließend auch die Erhebung der Daten durchgeführt. Zu diesem Zweck wurde für alle teilnehmenden Fellows ein individueller Zugang zur Online-Umfrage angelegt. Nach der Aktivierung der Online-Umgebung erhielten alle Fellows per Email eine Einladung mit dem Link zur ersten Erhebung (bzw. später zur vierten Erhebung). Der Zugang zur Erhebung war jeweils über die persönliche Teach First Emailadresse möglich. Der individuelle Zugang ermöglichte den teilnehmenden Fellows die Umfrage auch nach einer Unterbrechung zu einem späteren Zeitpunkt fortzuführen.

Vor Beginn der Umfrage wurde allen Teilnehmern die Anonymisierung ihrer Daten garantiert. Aus diesem Grund wurde die eingegebene Emailadresse nicht in den auszuwertenden Datenpool übernommen. Um dennoch eine Vergleichbarkeit der Daten zwischen den einzelnen Wellen zu ermöglichen, wurde in allen Erhebungen ein Code erstellt, der sich aus den Antworten auf vier Fragen zusammensetzt: Der Frage nach dem Tag der Geburt, dem zweiten Buchstaben des Vornamens der Mutter, dem letzten Buchstaben des Vornamens und dem ersten Buchstaben des Geburtsortes.

Anhand des auf diese Weise erstellten Codes konnte allen Fellows anonymisiert eine ID zugeteilt werden, welche die Zuordnung der Individualdaten über die vier Erhebungszeitpunkte möglich macht. Dadurch kann die individuelle Entwicklung der Fellows im Zeitraum der gesamten Teach First Vorbereitung verfolgt werden.

Die zweite und dritte Welle der Befragungen wurden vor Ort als Paper-Pencil-Befragungen durchgeführt. Für die zweite Welle wurde der auszufüllende Bogen mit Hilfe des Programms IBM SPSS Data Collection erstellt, um die Übernahme der erhobenen Daten später zu erleichtern und direkt an eine durch das Programm vorstrukturierte SPSS-Datei zu koppeln. Die Handhabung der Software stellte sich in Bezug auf Papierbefragungen allerdings als äußerst nutzerunfreundlich heraus. Aus diesem Grund wurde der Umfragebogen für die dritte Welle mit Libre Office erstellt. Die Daten wurden in diesem Fall anschließend in eine durch das Forschungsteam entsprechend der vorherigen Wellen strukturierte SPSS-Datei eingegeben.

### 5.1 Kontinuität in der Zusammensetzung der Stichprobe

Bei Erhebungen, die in mehreren Wellen die gleichen Personen befragen (Panel-Befragung), kommt es häufig vor, dass Personen nicht an allen Wellen der Befragung teilnehmen bzw. aus der Befragung ausscheiden – man spricht hier von Panel-Mortalität. Das hat zur Folge, dass die Daten einzelner Personen nicht für alle Erhebungszeitpunkte vorliegen. Typischerweise nimmt die Panel-Mortalität mit der Anzahl der Erhebungswellen zu. Abbildung 3 visualisiert, wie sich die TFD Stichprobe über die vier erhobenen Wellen zusammensetzt.

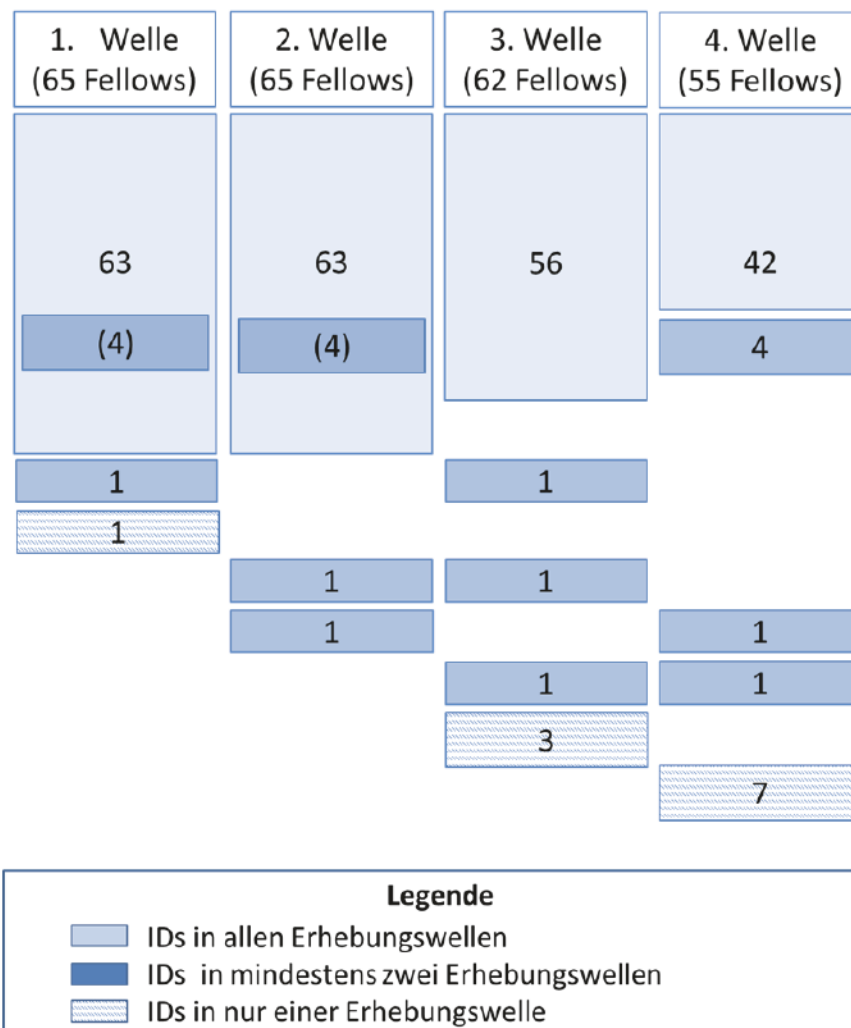


Abbildung 3: Zusammensetzung der Stichprobe über die vier Erhebungswellen

Aus Abbildung 3 ist zu entnehmen, dass an der ersten Befragung 65 Personen teilnahmen. Von diesen 65 Personen nahmen 63 auch an der 2. Befragung teil. Eine der beiden Personen (ID:41) hat nur an der ersten Befragung teilgenommen, die zweite Person (ID:44), hat an der 2. Befragung nicht teilgenommen, dafür aber wieder an der 3. Befragung. An der 2. Befragung nahmen insgesamt 65 Personen teil, d.h. zu den 63 Personen aus der ersten Welle kamen in der 2. Welle noch 2 neue Personen hinzu, die nicht an der ersten Befragung teilgenommen haben. Dies hat mit der Nachrekrutierung von TFD Fellows zu tun. An der 3. Befragung nahmen insgesamt 62 Personen teil. 57 Personen können aus der 1. Welle zugeordnet werden und ebenfalls 57 Personen können aus der 2. Welle zugeordnet werden. 3 Personen haben nur an den ersten beiden Wellen teilgenommen, nicht aber an der 3. und 4. Welle. 4 Personen sind in Welle 3 neu dazugekommen, die an den ersten beiden Erhebungen noch nicht teilgenommen hatten, eine Person ist in Welle 3 wieder dabei, die an Welle 2 nicht, aber an Welle 1 teilgenommen hatte (ID: 44 s.o.). Eine Person hat nur an Welle 2 und Welle 3 teilgenommen (ID: Neu2). In Welle 4 gibt es insgesamt 7 Personen, die anhand des individualisierten Codes keiner der vorherigen Wellen zugeordnet werden können. In der Einleitung des Fragebogens wurde explizit darauf hingewiesen, dass nur

Personen an der vierten Befragung teilnehmen mögen, die schon an einer der vorherigen Befragungen teilgenommen haben. Es handelt sich hier also entweder um Personen, die trotz des Hinweises gerne an der Befragung teilnehmen wollten oder um Personen, die ihren individuellen Code unabsichtlich oder absichtlich falsch eingegeben haben. Vier Personen haben nur an den ersten beiden und der vierten Welle teilgenommen (IDs: 3, 10, 17, 21), eine Person hat nur an Welle 2 und Welle 4 teilgenommen (ID: Neu1) und eine Person hat nur an den letzten beiden Wellen teilgenommen (ID: Neu4)<sup>2</sup>. Dadurch ergeben sich folgende Auswertungsperspektiven: 42 Personen haben an allen Befragungen teilgenommen. 46 Personen haben an Welle 1, 2 und 4 teilgenommen und können daher bezüglich des ausführlichen Tests verglichen werden. 56 Personen haben an den ersten drei Befragungen teilgenommen. Von 66 Personen liegen Daten zu mindestens zwei Messzeitpunkten vor und von 12 Personen wurden nur einmalig Daten erhoben. Dies hat zur Folge, dass den einzelnen Analysen teilweise unterschiedliche Stichprobengrößen zu Grunde liegen.

---

<sup>2</sup> Eine Übersicht der ID-Nummern der Personen, die nicht in allen Wellen enthalten sind findet sich in Tabelle 47 im Anhang.



## 6 Persönliche Voraussetzungen der Fellows

An der Evaluationsstudie nahmen ausschließlich TFD-Fellows teil, die im Jahr 2014 ihr Engagement begonnen haben. Die ersten Fellows erhielten ihre Zusage zur Teilnahme am TFD-Programm im August 2013, die letzten erhielten ihre Zusage im April 2014.

Die folgende Beschreibung von persönlichen Voraussetzungen der Fellows gründet im Wesentlichen auf der ersten Befragungswelle, da fast alle Personenmerkmale einmalig in dieser erhoben wurden.<sup>3</sup>

Die Geschlechterzusammensetzung der Stichprobe fällt zu Gunsten der Frauen aus: 60% der teilnehmenden Fellows sind weiblich. Das Durchschnittsalter der befragten Fellows beträgt 27,8 Jahre. Die jüngste Person ist 24 Jahre, die älteste Person 38 Jahre alt. Hauptsächlich handelt es sich bei den Befragten um Deutsch-Muttersprachler. Nur 17 % sprechen in ihrer Herkunftsfamilie zusätzlich noch eine andere Sprache und 2 Personen generell eine andere Sprache als Deutsch. Sowohl die Abiturnoten als auch die Noten der universitären Ausbildung sind überdurchschnittlich gut. Der Schnitt der Abiturnote liegt bei 1,9, der des ersten angegebenen Studienganges bei 1,6 und der des zweiten angegebenen Studienganges bei 1,4 Notenpunkten. Die Studiengänge werden in die Bereiche Sprach- und Kulturwissenschaften, Rechts-, Wirtschafts- und Sozialwissenschaften, Mathematik, Naturwissenschaften und Ingenieurwissenschaften kategorisiert. Eine genaue Aufschlüsselung der einzelnen Studiengänge innerhalb der Kategorie findet sich in Tabelle 46.

Am stärksten vertreten ist die Gruppe von Fellows, die Rechts-, Wirtschafts- und Sozialwissenschaften (46,8%). In etwa gleich groß ist die Gruppe der Fellows, die Sprach- und Kulturwissenschaften studiert haben (41,9%). Sechs Personen haben Mathematik oder Naturwissenschaften (9,7%) und nur eine Person hat ein Studium der Ingenieurwissenschaften absolviert (1,6%). 37 Personen haben einen zweiten Studiengang angegeben. Hier fällt die Verteilung zugunsten der Sprach- und Kulturwissenschaften aus (56,8%), die übrigen Fellows (43,2%) verteilen sich auf die Gruppe der Rechts-, Wirtschafts- und Sozialwissenschaften. Drei Personen haben einen dritten Studienabschluss angegeben und zwar alle einen Abschluss, der den Rechts-, Wirtschafts- und Sozialwissenschaften zuzuordnen ist.

In ihren pädagogischen Vorerfahrungen unterscheiden sich die Befragten sehr stark. 60% der Befragten haben bereits mit Kindern in einem schulischen Kontext gearbeitet. Die angegebene Dauer reichte von einem bis 72 Monaten. Mehr als die Hälfte dieser Personen (51,3%) gab eine Zeitdauer von 6 Monaten und mehr an. 81,5% der teilnehmenden Fellows haben bereits mit Kindern im außerschulischen Kontext gearbeitet. Hier variiert die Dauer noch stärker (Zeitspanne: 1 bis 130 Monate). 74,6% gaben eine Zeitdauer von einem Jahr

---

<sup>3</sup> Die Fragen nach dem Vorbereitungsdienst und Prüfungsverfahren für Lehrämter wurden in der zweiten Erhebung gestellt und das Belastungserleben wurde mehrfach erhoben.

und länger an. Keiner der Befragten hat jemals in einem Prüfungsverfahren oder Vorbereitungsdienst für Lehrämter gestanden. Insgesamt gaben 56 Personen an im Ausland studiert oder gearbeitet zu haben. 57,1% waren 12 Monate und länger im Ausland.

Um zu erfassen, welche Vorstellungen die Befragten zur Ausgestaltung ihrer Rolle als Fellow haben, wurden statt einfachen Einschätzungsitems paarweise Einschätzungsitems entwickelt (forced choice items). Diese Art der Präsentation soll verhindern, dass die Befragten allen Items stark zustimmen. Die beiden Items eines Paares stellen dabei zwei unterschiedliche Auffassungen dar, zwischen denen sich die Befragten auf einer vierstufigen Skala verorten müssen. Es darf also nur ein Kreuz gesetzt werden. Das Antwortformat lautet „Ich stimme dieser Ansicht völlig zu“ und „Ich stimme dieser Ansicht eher zu“. Die Mittelwerte sind hier so zu interpretieren, dass Werte die kleiner sind als 2,50 eine Tendenz zur Ansicht zeigen, die auf der linken Seite präsentiert wurde und Werte die größer sind als 2,50 eine Tendenz zur Ansicht zeigen, die auf der rechten Seite präsentiert wurde. Insgesamt liegen Angaben von 55 Fellows vor. Einige Fellows haben zu dieser Skala Kommentare abgegeben, aus denen ersichtlich wird, dass diese Art der Einschätzung teilweise zu einer Ablehnung geführt hat (z.B. „Aussagen sehr plakativ und zu wenig Möglichkeit zu Alternativen!“; „Alternative Aussagen wären schön.“).

Die Fellows, die Angaben zu ihrem Rollenverständnis machen, legen im Durchschnitt eher Wert darauf, mit einzelnen Schülern und Schülerinnen möglichst intensiv zu arbeiten als möglichst viele Schüler und Schülerinnen zu erreichen und gehen davon aus, dass ihre Schüler und Schülerinnen am meisten davon profitieren, wenn sie auf deren Stärken fokussieren, statt bei deren Defiziten anzusetzen (Mittelwerte: 2,90; 3,30). Sie beabsichtigen, sich an den Interessen der Schüler und Schülerinnen zu orientieren und deren Anliegen auch bei der Entwicklung neuer Konzepte zu berücksichtigen (Mittelwerte: 2,43; 1,70). Die Fellows möchten im Durchschnitt eher die Identität der Schüler und Schülerinnen gemeinsam weiterentwickeln und Schülern und Schülerinnen mit unterschiedlichem Migrationshintergrund Stolz auf ihre Herkunft vermitteln (Mittelwerte: 2,36; 3,27). Die Fellows haben eher vor, sich an den TFD Methoden zu orientieren als ein eigenes Profil an der Schule zu entwickeln (Mittelwert: 2,23). Den Fellows ist wichtig, dass die Schüler und Schülerinnen ihr Leistungspotential durch sie erkennen und Anerkennung erfahren für ihr Potential zur Veränderung (Mittelwerte: 1,60; 2,72).

	Rollenselbstbild der Fellows			
Variable	MZP (Welle 2)			
		M	STD	
	<b>Einleitungstext:</b> Die Rolle des Fellows kann unterschiedlich von Ihnen ausgeformt werden, im Folgenden interessiert und, wo Sie persönliche Schwerpunkte setzen.			
rolle1_2	In meiner Arbeit ist es mir wichtig, eine möglichst große Zahl von Schülern/Schülerinnen zu erreichen.	2,90	0,73	In meiner Arbeit ist es mir wichtig, mit einzelnen Schülern/Schülerinnen möglichst intensiv zu arbeiten.
rolle2_2	Meine Schüler/innen profitieren am meisten von meinem Einsatz, wenn ich an ihren Defiziten ansetze.	3,30	0,72	Meine Schüler/innen profitieren am meisten von meinem Einsatz, wenn ich auf ihre Stärken fokussiere.
rolle3_2	Meine Absicht ist, Schüler/innen mit unterschiedlichem Migrationshintergrund Stolz auf Ihre Herkunft zu vermitteln.	2,36	0,91	Meine Absicht ist, den Migrationshintergrund der unterschiedlichen Schüler/innen nicht selbst zum Thema zu machen.
rolle4_2	In meiner zukünftigen Arbeit orientiere ich mich vorrangig an den Methoden des Programms von Teach First.	2,23	0,65	In meiner zukünftigen Arbeit versuche ich vorrangig ein eigenes Profil in der Schule zu entwickeln.
rolle5_2	Entscheidendes Moment meiner Pädagogik wird sein, dass die Schüler/innen bei mir Anerkennung erfahren, so, wie sie sind.	2,72	1,00	Entscheidendes Moment meiner Arbeit wird sein, dass die Schüler/innen bei mir Anerkennung erfahren für ihr Potential zur Veränderung.
rolle6_2	Jeder Fellow sollte von den Anliegen der Schüler/innen ausgehen, wenn er neue Konzepte entwickelt.	1,70	0,63	Jeder Fellow kann etwas selbst am besten, darauf sollte er sich stützen, wenn er neue Konzepte entwickelt.
rolle7_2	Die Identität der Schüler/innen ist deren private Angelegenheit, die ich nicht zum Thema machen werde.	3,27	0,78	Im Kontakt mit mir werden die Schüler/innen ihre eigene Identität weiterentwickeln können.
rolle8_2	Mir ist wichtig, dass die Schüler/innen bei mir ihr Leistungspotential erkennen.	1,60	0,76	Mir ist wichtig, dass die Schüler/innen von mir Solidarität mit ihrer Ausgangslage erfahren.
rolle9_2	Ich werde mich weitgehend an den Interessen der Schüler/innen orientieren.	2,43	0,79	Ich werde versuchen, den Schülerinnen neue Interessen nahe zu bringen.
N <sub>Max</sub>		55		

Tabelle 2: Rollenselbstbild der Fellows - Deskriptive Statistiken

Im Gegensatz zu den persönlichen Hintergrundmerkmalen, die nur einmal erhoben wurden, weil sie sich in der Regel nicht verändern, wurde das Belastungserleben der Fellows in Erhebungswelle 2, 3 und 4 erfasst. Die Fellows sollten auf einer Skala von 1 – *gar nicht belastet*, bis 4 – *stark belastet* angeben, wie sehr sie sich durch verschiedene Faktoren belastet fühlen. Man erkennt eine Steigerung der erlebten Belastung vom ersten zum zweiten Messzeitpunkt. Besonders stark ist diese bei der Belastung durch die Menge der täglichen Arbeit zu erkennen. Dieser Belastungsfaktor hat zum zweiten Messzeitpunkt



seinen Höhepunkt (Mittelwert: 3,19). Während die anderen Belastungsfaktoren vom zweiten zum dritten Messzeitpunkt entweder gleich belastend oder als belastender eingeschätzt werden, wird die Menge an täglicher Arbeit wieder als etwas weniger belastend empfunden. Als am stärksten belastend wird zum dritten Messzeitpunkt der kurzfristige Zeitdruck wahrgenommen (Mittelwert: 3,05).

Variable	Belastungserleben	MZP 1 (Welle 2)		MZP 2 (Welle 3)		MZP 3 (Welle 4)	
	<b>Einleitungstext:</b> Wie sehr fühlen Sie sich in den letzten Wochen belastet durch...						
		<b>M</b>	<b>STD</b>	<b>M</b>	<b>STD</b>	<b>M</b>	<b>STD</b>
gstr2_2/ gstr2_3/ gstr2_4	... die Menge ihrer täglichen Arbeit?	2,74	1,05	3,19	0,79	2,85	0,77
gstr3_2/ gstr3_3/ gstr3_4	... kurzfristigen Zeitdruck?	2,74	0,97	3,05	0,82	3,04	0,68
gstr5_2/ gstr5_3/ gstr5_4	... negative Ergebnisse Ihrer Bemühungen?	1,69	0,83	1,87	0,87	2,38	0,73
gstr6_3/ gstr6_4	... die Verantwortung für andere Menschen im Rahmen Ihrer Tätigkeit?	-	-	1,98	0,91	2,40	0,96
gstr7_2/ gstr7_3/ gstr7_4	... die Bewertung Ihrer Arbeit?	1,61	0,75	1,69	0,71	2,08	0,65
<b>N</b>		<b>65</b>		<b>62</b>		<b>48</b>	

Tabelle 3: Belastungserleben - Deskriptive Statistiken

Die Faktorenanalyse zum ersten Messzeitpunkt ergibt eine zweifaktorielle Lösung mit jeweils zwei Items pro Faktor, sodass hier eine Skalenbildung nicht sinnvoll ist. Zum ersten Messzeitpunkt wurde das Item gstr6 noch nicht miterhoben, weil die Fellows zu diesem Zeitpunkt noch keinen Kontakt zu Schülern und Schülerinnen hatten. Die Faktorenanalyse zum zweiten Messzeitpunkt ergibt eine zweifaktorielle Lösung. Die Items gstr5– gstr7 laden auf den ersten Faktor und die Items gstr2 und gstr3 laden auf den zweiten Faktor. Allerdings ergibt sich in der Reliabilitätsanalyse der Items des ersten Faktors keine zufriedenstellende interne Konsistenz (Cronbach's alpha: 0,69), sodass hier keine Skalenbildung möglich ist. Die Faktorenanalyse zum dritten Messzeitpunkt ergibt eine zweifaktorielle Lösung. Dabei laden die Items gstr3, gstr5 und gstr6 auf den ersten Faktor und die Items gstr2, gstr3 und gstr7 auf den zweiten Faktor. Allerdings weist das Item gstr3 praktisch identische Ladungen auf dem ersten und zweiten Faktor auf, sodass man es nicht eindeutig einem Faktor zuordnen kann. Dadurch laden auf jedem Faktor nur jeweils zwei Items, was für eine Skalenbildung nicht ausreichend ist.

## 7 Prozesse des Kompetenzerwerbs

Entsprechend des theoretischen Wirkmodells können Prozesse des Kompetenzerwerbs in Prozesse auf der Mesoebene und Prozesse auf der Mikroebene unterschieden werden. Auf der Mesoebene betrachten wir verschiedene Merkmale der Lernumgebung wie die Einschätzung der Qualität einzelner Trainingsteile, die Qualität der Lehrmaterialien oder die Kohärenz des Trainings. Auf der Mikroebene ist die Nutzung der Lernumgebung durch die Fellows verortet. Hier betrachten wir die Nutzung von Lernstrategien, von Lernzeit sowie von professioneller Kooperation. Im Folgenden wird zunächst auf die Merkmale der Lernumgebung und anschließend auf die Merkmale der Nutzung der Lernumgebung eingegangen.

### 7.1 Merkmale der Lernumgebung

Merkmale der Lernumgebung wurden in der zweiten, dritten und vierten Befragung erhoben. Während die zweite Erhebung auf die Lernumgebung Online Campus fokussierte, wurden in der dritten Befragung die Lernumgebungen Seminarteil, Känguru-Woche und MSact! evaluiert. In der vierten Befragung wurden die schuleinsatzbegleitenden Qualifizierungsmaßnahmen von TFD und Merkmale der Einsatzschule als neuer Lernumgebung für die Fellows erhoben.

#### Zweite Erhebung: Einschätzung der Qualität des Online Campus

In der zweiten Erhebung wurde die Lernumgebung Online Campus, eine online Lernplattform, auf der für jede Woche die Lernmaterialien (Reader und Reflexionsaufgaben) bereitgestellt werden, evaluiert. Die Plattform wird weiterhin von den Fellows genutzt, um Hausaufgaben hochzuladen und Feedback zu geben bzw. erhalten.

Um die Qualität des Online Campus zu erfassen, wurden zwei Instrumente entwickelt, welche die Gestaltungsqualität der Lernplattform und der Interaktion zwischen den Beteiligten über die Lernplattform abbilden sollten, zum anderen Instrumente, die die Qualität der Lehrmaterialien erfassen sollten. Zur Gestaltungsqualität der Lernplattform wurden Items zur Nutzerfreundlichkeit sowie zur Struktur der Lernplattform, zur Handhabung der Tools auf der Lernplattform und zum technischen Support entwickelt. In Bezug auf die Interaktion zwischen den Beteiligten wurde sowohl die Interaktion zwischen den Fellows als auch die Interaktion zwischen Fellows und Trainern / Trainerinnen berücksichtigt. Um die Qualität der Lehrmaterialien zu bewerten, wurden Items zu den Aspekten Menge des Inputs, Strukturiertheit, Verständlichkeit, Praxisorientierung, wissenschaftliche Fundiertheit und kognitive Aktivierung entwickelt. Die Items waren alle auf einer Skala von 1 – *trifft nicht zu* bis 4 – *trifft zu* einzuschätzen.

##### 7.1.1 Gestaltungsqualität der Lernplattform

In Bezug auf Nutzerfreundlichkeit erhält die Lernplattform sehr positive Einschätzungen von den Befragten. Alle Mittelwerte liegen deutlich über 3,00, d.h. die Mehrheit der Befragten hat bei diesen Items *trifft eher zu* und *trifft zu* angegeben. Die Befragten gaben an, dass die Plattform selbsterklärend gestaltet und die Anleitung zur Nutzung hinreichend ist sowie dass

sie wissen, an wen sie sich bei auftretenden Problemen wenden können. Auch die Navigation auf den Seiten des Online Campus wird als gut funktionierend bewertet. Ähnlich verhält es sich mit der Struktur der Lernplattform und der Handhabung der Tools. Besonders die Ordnerstruktur wird als übersichtlich bewertet (Mittelwert: 3,58), während die Übersichtlichkeit der Foren nicht ganz so gut (Mittelwert: 2,69) eingeschätzt wird. Die Benutzung der Tools (Forenbeiträge schreiben, Dokumente hochladen) sowie der technische Support werden als durchgängig positiv eingeschätzt (Mittelwerte: 3,53-3,73). Tabelle 4 enthält die deskriptiven Statistiken der Einzelitems.

Variable	Nutzerfreundlichkeit	MZP (Welle 2)	
		M	STD
tec_oc1_2	Die Navigation auf den Seiten des Online Campus funktioniert gut.	3,34	0,57
tec_oc2_2	Die Lernplattform Online Campus ist selbsterklärend gestaltet.	3,44	0,64
tec_oc3_2	Die Anleitung zum Online Campus ist hinreichend.	3,77	0,43
soc_oc1_2	Wenn ich Probleme mit der Nutzung des Online Campus habe, weiß ich an wen ich mich wenden kann.	3,92	0,27
	<b>Struktur</b>	<b>M</b>	<b>STD</b>
tec_oc4_2	Die Ordnerstruktur der Lernplattform Online Campus ist übersichtlich gestaltet.	3,58	0,53
tec_oc5_2	Die Forenstruktur der Lernplattform Online Campus ist übersichtlich gestaltet.	2,69	0,75
	<b>Handhabung der Tools</b>	<b>M</b>	<b>STD</b>
tec_oc6_2	Es ist technisch einfach einen Beitrag im Forum zu schreiben.	3,53	0,59
tec_oc7_2	Es ist technisch einfach ein Dokument auf der Lernplattform hochzuladen	3,63	0,55
soc_oc2_2	Wenn ich Probleme mit der Nutzung des Online Campus habe, wird mir schnell geholfen.	3,73	0,51
N		65	

Tabelle 4: Gestaltungsqualität Lernplattform - Deskriptive Statistiken der Einzelitems

Die Faktorenanalyse zeigt, dass die unterschiedlichen Aspekte der Gestaltungsqualität im Wesentlichen auf je eigene Faktoren laden. Lediglich die beiden Items zum technischen Support bilden keinen eigenen Faktor. Das erste Item lädt auf den Faktor zur Nutzerfreundlichkeit, das andere auf den Faktor zur Handhabung der Tools (mit negativer Ladung). Einzelne Skalen können hier nicht berechnet werden, da Reliabilitätsanalysen zeigen, dass die interne Konsistenz nicht zufriedenstellend ist.

### 7.1.2 Interaktion zwischen den Fellows und den Trainern / Trainerinnen

Um die Interaktion zwischen den Beteiligten abzubilden, wurden 9 Items zur sozialen Eingebundenheit entwickelt, die sowohl die Interaktion zwischen den Fellows als auch zwischen Fellows und Trainern / Trainerinnen abbilden. Insgesamt wurden die Items zur Interaktion zwischen Fellows und Trainern schlechter bewertet (Mittelwerte: 1,91-2,70) als die Items zur Interaktion zwischen den Fellows (Mittelwerte: 2,61-3,34). Im Durchschnitt bewerten die Befragten den Austausch mit den Trainern und Trainerinnen als eher nicht ausreichend, die Rückmeldungen der Trainer als eher nicht hilfreich, dafür werden sie aber eher



als zeitnah erfolgend eingeschätzt. Dagegen wird der Austausch mit anderen Fellows durchschnittlich als eher gut organisiert sowie ausreichend bewertet. Auch die Rückmeldungen von anderen Fellows werden im Mittel als eher gewinnbringend eingeschätzt. Tabelle 5 zeigt die Statistiken der Einzelitems zur Interaktion.

Variable	Interaktion zwischen Fellows und Trainern	MZP (Welle 2)	
		M	STD
soc_oc5_2	Der Austausch mit den Trainern findet für mich in ausreichendem Maße statt.	2,40	0,98
soc_oc7_2	Die Rückmeldungen der Trainer, die ich zu meinen Beiträgen erhalten habe, finde ich hilfreich.	1,91	0,94
soc_oc9_2	Die Rückmeldungen, die ich zu meinen Beiträgen erhalte, erfolgen zeitnah.	2,70	0,89
Variable	Interaktion zwischen Fellows	M	STD
soc_oc3_2	Der Austausch mit anderen Fellows ist gut organisiert.	3,12	0,76
soc_oc4_2	Der Austausch mit anderen Fellows findet für mich in ausreichendem Maße statt.	3,34	0,83
soc_oc6_2	Die Rückmeldungen von anderen Fellows, die ich zu meinen Beiträgen bekomme, sind gewinnbringend.	2,61	0,75
N		65	

Tabelle 5: Interaktion zwischen den Beteiligten – Deskriptive Statistiken

Von den ursprünglich 9 Items zur sozialen Eingebundenheit wurde nach Faktoren- und Reliabilitätsanalysen je eine Skala mit 3 Items zu Interaktion zwischen Fellows und Interaktion zwischen Fellows und Trainern bzw. Trainerinnen gebildet. Der Mittelwert der Skala *Interaktion zwischen Trainern bzw. Trainerinnen und Fellows* beträgt 2,30 und liegt etwas unterhalb des theoretischen Mittelwertes. Der Mittelwert der Skala *Interaktion zwischen Fellows* liegt mit 3,01 über dem theoretischen Mittelwert. Die Interaktion zwischen Fellows wird demnach als eher positiv eingeschätzt, wie bereits bei der Erläuterung der Einzelitems deutlich wurde. Zusätzlich wurde gefragt, wie häufig die teilnehmenden Fellows über die Anwendungen im Online Campus Kontakt zueinander haben. Die Antwortmöglichkeiten lauteten: *Mehrmals am Tag, einmal am Tag, mehrmals pro Woche, einmal pro Woche, weniger als einmal pro Woche*. Die Auswertung ergibt, dass die meisten der Befragten (46,9%) mehrmals pro Woche Kontakt über die Plattform Online Campus haben. Mehr als die Hälfte der teilnehmenden Fellows (56,3%) haben mehrmals pro Woche oder häufiger Kontakt zueinander (siehe Tabelle 45).

### 7.1.3 Qualität der Lehrmaterialien

Insbesondere die Menge an Inhalten, die Strukturiertheit und Verständlichkeit der Lehrmaterialien wird von den Befragten sehr positiv beurteilt. Alle Mittelwerte liegen deutlich über 3,00, d.h. die Mehrheit der Befragten hat bei diesen Items *trifft eher zu* und *trifft zu*

angegeben.<sup>4</sup> Die Menge an Input wurde im Durchschnitt als angemessen beurteilt. Dies trifft sowohl auf die Menge der Inhalte, die jede Woche zu bearbeiten sind, als auch auf die Portfolioaufgaben zu. Die Reader werden nicht als inhaltlich überladen wahrgenommen.

Variable	Menge des Inputs	MZP (Welle 2)	
		M	STD
inp_oc1_2	Die Menge der Inhalte pro Woche ist gut zu verarbeiten.	3,12	0,78
inp_oc5_2	Die Anzahl der Portfolioaufgaben pro Woche finde ich angemessen.	3,22	0,76
inp_oc9_2	Die einzelnen Reader sind inhaltlich überfrachtet.	1,55	0,69
Variable	Strukturiertheit	M	STD
strukt_oc2_2	Innerhalb der Materialien ist ein roter Faden für mich erkennbar.	3,54	0,53
strukt_oc6_2	Innerhalb der Materialien werden die wesentlichen Punkte hervorgehoben.	3,34	0,67
strukt_oc10_2	Innerhalb der Materialien werden die Inhalte willkürlich nacheinander präsentiert.	1,55	0,61
strukt_oc13_2	Die einzelnen Fächer sind sinnvoll aufeinander bezogen.	3,25	0,61
strukt_oc15_2	Beim Lesen der Readertexte bin ich unsicher darüber, was die Lernziele sind.	1,62	0,65
Variable	Verständlichkeit	M	STD
verst_oc3_2	Der Text der Reader des Online Campus ist in einer verständlichen Sprache geschrieben.	3,88	0,33
verst_oc7_2	Es macht mir manchmal Mühe die Inhalte nachzuvollziehen.	1,82	0,81
verst_oc11_2	In den Readern des Online Campus wird mit anschaulichen Beispielen gearbeitet.	3,52	0,56
N		65	

Tabelle 6: Qualität der Lehrmaterialien: Menge des Inputs, Strukturiertheit, Verständlichkeit – Deskriptive Statistiken

Die Materialien werden als strukturiert und verständlich beurteilt. Die teilnehmenden Fellows geben an, dass innerhalb der Materialien ein roter Faden erkennbar ist, die einzelnen Fächer aufeinander bezogen sind und wesentliche Punkte hervorgehoben werden. Inhalte wirkten nicht willkürlich nacheinander präsentiert und auch die Lernziele würden deutlich. Die Befragten beurteilen den Text der Reader in großem Maße als verständlich (Mittelwert: 3,88) und geben an, dass mit anschaulichen Beispielen gearbeitet wird. Es macht ihnen keine Mühe, die Inhalte nachzuvollziehen.

Auch die Praxisorientierung des Materials wird eher positiv bewertet. Hier liegen die Mittelwerte alle um 3,00 – insgesamt etwas niedriger als bei den zuvor beschriebenen Aspekten. Am stärksten zugestimmt wurde der Aussage, dass die Inhalte so aufbereitet sind, dass der Praxisbezug deutlich wird (Mittelwert 3,38). Im Durchschnitt gaben die Befragten weiter an, dass die Portfolioaufgaben eher hilfreich für die Vorbereitung auf die Schulpraxis sind und dass sie durch die Bearbeitung der Reader notwendiges Handwerkszeug für die Praxis erhalten.

<sup>4</sup> Hierbei ist zu beachten, dass es teilweise Items mit Negativaussagen gab. Die Mittelwerte dieser Items liegen alle unter 2,0, d.h. hier haben die meisten Fellows *trifft nicht zu* und *trifft eher nicht zu* angegeben.



Variable	Praxisorientierung	MZIP (Welle 2)	
		M	STD
prax_oc4_2	Die Portfolioaufgaben helfen mir, mich auf die Praxis in der Schule vorzubereiten.	3,05	0,69
prax_oc8_2	Durch die Bearbeitung der Reader erhalte ich das notwendige Handwerkszeug für die Praxis.	2,89	0,61
prax_oc12_2	Die Inhalte der Reader sind so aufbereitet, dass der Bezug zur Praxis deutlich wird.	3,38	0,68
N		65	

Tabelle 7: Qualität der Lehrmaterialien: Praxisorientierung – Deskriptive Statistiken

Die teilnehmenden Fellows nehmen das Lehrmaterial als eher kognitiv aktivierend wahr. So stimmen sie eher zu, dass das Material zum Nachdenken anregt, abwechslungsreich gestaltet ist, dazu anregt, Verbindungen zu früheren Erfahrungen herzustellen und neue Themen selbstständig zu erkunden. Zudem stimmen die Fellows Aussagen zu, dass das Material Beispiele enthält, die zu tieferem Verständnis beitragen und dazu anregt, sich mit Aufgaben oder Themen eingehender zu beschäftigen. Die deskriptiven Statistiken der Einzelitems werden in Tabelle 8 berichtet.

Variable	Kognitive Aktivierung	MZIP (Welle 2)	
		M	STD
kogaktiv_oc1_2	Das Material regt mich an, Verbindungen zu früheren Erfahrungen herzustellen.	3,27	0,70
kogaktiv_oc2_2	Das Material enthält Beispiele, die zu einem tiefergehenden Verständnis beitragen.	3,36	0,65
kogaktiv_oc3_2	Das Material regt mich zum Nachdenken an.	3,38	0,68
kogaktiv_oc4_2	Das Material ist abwechslungsreich gestaltet.	3,30	0,71
kogaktiv_oc5_2	Das Material regt mich dazu an, mich mit interessanten Aufgaben oder Themen eingehender zu beschäftigen.	3,00	0,71
kogaktiv_oc6_2	Das Material ist so gestaltet, dass ich neue Themen selbstständig erkunden kann.	3,03	0,71
kogaktiv_oc7_2	Das Material bringt mich dazu neue Theorien an Beispielen aus dem täglichen Leben auszuprobieren.	2,98	0,88
N		65	

Tabelle 8: Qualität der Lehrmaterialien: Kognitive Aktivierung – Deskriptive Statistiken

Lediglich die wissenschaftliche Fundierung des Lehrmaterials wird etwas schlechter beurteilt. Die Aussage, dass in den Readern ausreichend wissenschaftliche Studien vorgestellt werden, wird mit einem Mittelwert von 2,37 von den Fellows am niedrigsten eingeschätzt, während sie die Bezugnahme auf aktuelle Forschungsergebnisse mit einem Mittelwert von 2,84 am höchsten einschätzen. Die Mittelwerte aller Einschätzungen zu wissenschaftlicher Fundierung werden in Tabelle 9 berichtet.

Variable	Wissenschaftliche Fundiertheit	MZP (Welle 2)	
		M	STD
wissfund_oc1_2	In den Readern, die mir auf dem Online Campus zur Verfügung gestellt werden, wird Bezug auf aktuelle Forschungsergebnisse genommen.	2,84	0,72
wissfund_oc2_2	In den Readern, die mir auf dem Online Campus zur Verfügung gestellt werden, werden Auszüge aus wissenschaftlichen Artikeln zitiert.	2,78	0,83
wissfund_oc4_2	In den Readern, die mir auf dem Online Campus zur Verfügung gestellt werden, werden ausreichend wissenschaftliche Studien aus Pädagogik und/oder Psychologie vorgestellt.	2,37	0,66
wissfund_oc5_2	In den Readern, die mir auf dem Online Campus zur Verfügung gestellt werden, wird ausreichend auf wissenschaftliche Theorien Bezug genommen.	2,53	0,69
N		65	

Tabelle 9: Qualität der Lehrmaterialien: Wissenschaftliche Fundierung – Deskriptive Statistiken

Die fünf Items, die entwickelt wurden, um Strukturiertheit abzubilden, laden in der Faktorenanalyse alle auf einen Faktor. In der Reliabilitätsanalyse wird für Cronbach's alpha ein Wert von 0,77 ausgegeben, was für eine akzeptable Reliabilität spricht. Auch die Items, die zu den Aspekten Menge an Input und Praxisorientierung entwickelt wurden, laden jeweils auf einen Faktor. Allerdings können hier keine zufriedenstellenden Werte für Cronbach's alpha erzielt werden. Aus diesem Grund werden aus diesen Items keine Skalen gebildet. Für die Items zum Aspekt Verständlichkeit ergibt die Faktorenanalyse eine zweifaktorielle Lösung mit jeweils zwei Items auf jedem Faktor. Daher kann auch hier keine Skala gebildet werden. Zur wissenschaftlichen Fundierung wurden insgesamt 6 Items gebildet. Die Faktorenanalyse zeigt, dass die ersten fünf Items auf den ersten Faktor laden und das sechste alleine einen zweiten Faktor bildet. Daher wird das sechste Item verworfen. Die Reliabilitätsanalyse ergibt unter Ausschluss von wissfund\_oc3 ein Cronbach's alpha von 0,75, was für eine akzeptable Reliabilität spricht. Der Mittelwert der neu gebildeten Skala *wissenschaftliche Fundierung* beträgt 2,50. Die Items, die zur kognitiven Aktivierung entwickelt wurden, laden in der Faktorenanalyse alle auf einen Faktor. Die anschließende Reliabilitätsanalyse ergibt ein Cronbach's alpha von 0,83, was für eine zufriedenstellende Reliabilität spricht. Der Mittelwert der Skala *kognitive Aktivierung* beträgt 3,15.

### Dritte Erhebung: Einschätzung der Praxisteile der Sommerakademie

In der dritten Erhebungswelle wurden die Erfahrungen der Praxisteile der Sommerakademie evaluiert. Dazu zählen der Seminarteil auf Burg Katlenburg, der die Inhalte aus dem Online Campus vertiefen soll, die Känguru-Woche, bei der ein erfahrener Fellow (als Känguru-Mutter) einen neuen Fellow (Känguru-Baby) eine Woche lang mit an seine Einsatzschule nimmt und auf den bevorstehenden Schuleinsatz einstimmt, und schließlich MSact!, ein Sommerschulprogramm für Schüler an Brennpunktschulen, bei dem die Fellows erste Erfahrungen im Unterrichten und in der Projektarbeit mit Schülern sammeln sollen.

#### 7.1.4 Seminarteil Katlenburg

Um die Erfahrungen mit dem Seminarteil zu erfassen, wurden sowohl Items zur Kooperation zwischen den Fellows (6 Items) als auch Items zur Gestaltung der Lernumgebung "Seminarteil Katlenburg" (5 Items) entwickelt. Die Befragten schätzen die Kooperation untereinander sehr hoch ein. Die Mittelwerte der Einzelitems liegen fast alle um 3,50 herum. Besonders hoch wurde die gegenseitige Unterstützung beim Lernen eingeschätzt (Mittelwert: 3,70).

Variable	Professionelle Kooperation zwischen den Fellows	MZP (Welle 3)	
		M	STD
erf_kat2_3	Während des Seminarteils auf der Katlenburg haben sich die Fellows gegenseitig beim Lernen unterstützt.	3,70	0,46
erf_kat3_3	Während des Seminarteils auf der Katlenburg entstanden durch Fragen aus der Gruppe neue Lösungsansätze für Schwierigkeiten in Unterrichtssituationen.	3,43	0,67
erf_kat5_3	Während des Seminarteils auf der Katlenburg haben sich die Fellows gegenseitig Feedback zur Entwicklung ihrer pädagogischen Kompetenzen gegeben.	3,20	0,81
erf_kat9_3	Während des Seminarteils auf der Katlenburg habe ich mit anderen Fellows gemeinsam Fragen pädagogischen Handelns diskutiert.	3,48	0,62
erf_kat10_3	Während des Seminarteils auf der Katlenburg habe ich mich mit anderen Fellows über pädagogische Handlungsalternativen ausgetauscht.	3,33	0,81
erf_kat11_3	Während des Seminarteils auf der Katlenburg habe ich mit anderen Fellows über Lösungen in schwierigen Unterrichtssituationen beratschlagt.	3,49	0,70
N		62	

Tabelle 10: Professionelle Kooperation der Fellows während des Seminarteils auf der Katlenburg - Deskriptive Statistiken

Auch die Gestaltung der Lernumgebung wird von den befragten Fellows sehr positiv eingeschätzt. Sie gaben an, dass ihnen während des Seminarteils die Bedeutsamkeit der Inhalte des Online Campus eher deutlich wurde. Weiter stimmten sie eher zu, dass das im Online Campus erworbene Wissen vertieft und an praktischen Beispielen geübt wurde, um Handlungssicherheit zu gewinnen. Das einzige Item, dessen Mittelwert unter 3,00 liegt, bezieht sich auf die Besprechung von Fragen, die sich aus der Bearbeitung der Reader im Online Campus ergaben. Bei dieser Frage ist es möglich, dass nicht nur Personen wenig zustimmen, die das Gefühl hatten, dass ihre Fragen nicht besprochen wurden, sondern auch Personen wenig zustimmen, die keine Fragen hatten.

Variable	Gestaltung der Lernumgebung „Seminarteil Katlenburg“	MZP (Welle 3)	
		M	STD
erf_kat1_3	Während des Seminarteils auf der Katlenburg wurde die Bedeutsamkeit der Inhalte des Online Campus für den Schuleinsatz offensichtlich.	3,33	0,63
erf_kat4_3	Während des Seminarteils auf der Katlenburg konnte ich das Gelernte aus dem Online Campus an praktischen Unterrichtsbeispielen üben.	3,13	0,72
erf_kat6_3	Während des Seminarteils auf der Katlenburg wurde das Wissen aus dem Online Campus vertieft.	3,20	0,77
erf_kat7_3	Während des Seminarteils auf der Katlenburg wurden meine Fragen, die sich aus den Readern des Online Campus ergaben, besprochen.	2,95	0,87
erf_kat8_3	Während des Seminarteils auf der Katlenburg konnte ich durch konkrete Übungen Handlungssicherheit für Unterrichtssituationen gewinnen.	3,08	0,86
N		62	

Tabelle 11: Gestaltung des Trainingsteils "Seminarteil auf der Katlenburg" - Deskriptive Statistiken

Die Faktorenanalyse aller Items zu den Erfahrungen mit dem Seminarteil ergibt eine zweifaktorielle Lösung. Die Items zur Kooperation zwischen den Fellows laden auf den ersten Faktor, die Items zur Gestaltung der Lernumgebung auf den zweiten Faktor. Die Reliabilitätsanalysen ergeben für den ersten Faktor eine zufriedenstellende Reliabilität (Cronbach's alpha: 0,82) und für den zweiten Faktor eine akzeptable Reliabilität (Cronbach's alpha: 0,77). Der Mittelwert der Skala *Kooperation zwischen den Fellows* beträgt 3,44. Der Mittelwert der Skala *Gestaltung der Lernumgebung* beträgt 3,12.

#### 7.1.5 Känguru Erfahrung

Insgesamt wurden 15 Items zur Evaluierung des Trainingsteils „Känguru“ entwickelt, die auf die von TFD formulierten Lernziele für die Känguru-Woche abgestimmt sind. Dabei geht es im Wesentlichen um das Einfühlen in eine Schulkultur, das erste Einüben von Unterrichtshandlungen und Reflexionsarbeit zwischen zukünftigen Fellows und den Fellows, die bereits im Einsatz sind. Zusätzlich wurden Items entwickelt, welche die Nutzbarkeit von Wissen aus zuvor absolvierten Trainingsteilen in den darauffolgenden Trainingsteilen erfassen.

Zur Känguru-Woche ist zunächst anzumerken, dass von den 62 Befragten nur 50 Personen an der Känguru-Woche teilgenommen haben. Des Weiteren wird von einigen Fellows auf dem Fragebogen angemerkt, dass die Känguru-Woche kurz vor den Sommerferien stattfand, um zu begründen, warum den Aussagen nicht oder nur wenig zugestimmt wird. Dies spiegelt sich auch in den Mittelwerten vieler Items wider (Mittelwerte unter 2,50). Im Durchschnitt werden die Aussagen mit *trifft eher nicht zu* eingeschätzt. Besonders niedrig zugestimmt wird den Aussagen, dass die Möglichkeit bestand den Lernfortschritt von Schülern / Schülerinnen zu diagnostizieren (Mittelwert: 1,55) und dass die Gelegenheit bestand, mit dem erfahrenen Fellow über die eigenen Entwicklungsfelder zu reflektieren (Mittelwert: 1,92). Am höchsten stimmen die teilnehmenden Fellows den Aussagen zu, dass Gelegenheit



bestand, Einblick in die Schulkultur und laufende Schulprojekte zu bekommen (Mittelwerte: 3,12; 3,04).

Variable	Erfahrungen während der Känguru-Woche	MZP (Welle 3)	
		M	STD
erf_käng1_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, Einblick in die alltäglichen Unterrichtsabläufe zu bekommen.	2,67	1,05
erf_käng2_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, das bisher Gelernte in konkreten Situationen mit SchülerInnen umzusetzen.	2,37	1,01
erf_käng3_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, mein Wissen aus dem Seminarteil auf der Katlenburg zu vertiefen.	2,16	0,90
erf_käng4_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, das Gelernte aus dem Online Campus an praktischen Unterrichtsbeispielen zu üben.	2,16	0,90
erf_käng5_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, meine pädagogischen Fähigkeiten zu testen.	2,86	0,89
erf_käng6_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, Gruppen anzuleiten.	2,67	1,01
erf_käng7_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, Einblick in laufende Schülerprojekte zu bekommen.	3,04	1,04
erf_käng8_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, eine Unterrichtsstunde zu planen.	2,08	1,24
erf_käng9_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, von meinem Känguru Feedback zu meinem Unterrichtshandeln zu bekommen.	2,33	1,20
erf_käng10_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, mit meinem Känguru über meine Entwicklungsfelder zu reflektieren.	1,92	1,02
erf_käng11_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, den Lernfortschritt von SchülerInnen zu diagnostizieren.	1,55	0,84
erf_käng12_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, Einblicke in die Schulkultur der besuchten Schule zu bekommen.	3,12	0,83
erf_käng13_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, ein erstes Gefühl für das eigene Unterrichten bekommen.	2,24	1,03
erf_käng14_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, im Gespräch mit dem Känguru Lösungen für Schwierigkeiten in Unterrichtssituationen entwickeln.	2,35	0,97
erf_käng15_3	Während der „Känguru-Woche“ hatte ich Gelegenheit, im Gespräch mit dem Känguru Fragen, die das pädagogische Handeln betreffen, zu klären.	2,73	0,88
N		50	

Tabelle 12: Erfahrungen während des Trainingsteils "Känguru" - Deskriptive Statistiken

Die Faktorenanalyse ergibt eine vierfaktorielle Lösung, wobei auf den vierten Faktor nur zwei Items laden. Der Screeplot unterstreicht eine dreifaktorielle Lösung. Auf den ersten Faktor laden die Items erf\_käng8, erf\_käng9, erf\_käng10 und erf\_käng13. Sie bilden *Unterrichtsplanung / Unterrichtsentwicklung* (TFD: Unterrichten - Planen) ab. Die Reliabilitätsanalyse ergibt für diese Skala ein Cronbach's alpha von 0,89. Der Mittelwert für diese neu gebildete Skala beträgt 2,14. Auf den zweiten Faktor laden die Items erf\_käng3,

erf\_käng4, erf\_käng11 und erf\_käng14. Die Reliabilitätsanalyse zeigt, dass Item erf\_käng7 eine geringe Trennschärfe aufweist und auch Cronbach's alpha bei Ausschluss dieses Items anstiege. Aus diesem Grund wird die Skala ohne erf\_käng7 gebildet. Cronbach's alpha beträgt 0,84 und spricht für eine gute Reliabilität. Inhaltlich bilden die verbleibenden Items die Inhaltsdimension *Wissensanwendung / Theorie-Praxis Transfer* ab. Der Mittelwert dieser Skala beträgt 2,06. Auf den dritten Faktor laden die Items erf\_käng1, erf\_käng2, erf\_käng5 und erf\_käng6. Diese Items bilden *Unterrichtserfahrung* ab (TFD: Unterrichten- Umsetzen). Die Reliabilitätsanalyse ergibt für diese Items ein Cronbach's alpha von 0,82, welches sich durch Ausschluss von erf\_käng1 auf 0,85 verbessern lässt. Die auf diese Weise gebildete Skala hat einen Mittelwert von 2,63.

#### 7.1.6 MSact!

Zur Evaluierung des Trainingsteils MSact! wurden 9 Items entwickelt. An der Sommerschule MSact! nahmen 54 Fellows teil, die 8 Fellows, die auf den Schuleinsatz an einer Grundschule vorbereitet werden, nahmen am Ausbildungsteil CLIMB teil.

Variable	Erfahrungen mit MSact!	MZP (Welle3)	
		M	STD
erf_MSA1_3	Im Rahmen von MSact! konnte ich erkennen, dass ich viel mehr pädagogisches Wissen benötige.	2,48	0,87
erf_MSA2_3	Im Rahmen von MSact! konnte ich gute Beispiele für meinen Einsatz in der künftigen Schule erleben.	3,45	0,70
erf_MSA3_3	Im Rahmen von MSact! konnte ich die Anwendung diagnostischer Verfahren selbst einüben.	2,58	0,89
erf_MSA4_3	Im Rahmen von MSact! konnte ich lernen, wie ich Lerngruppen organisiere und manage.	3,58	0,65
erf_MSA5_3	Im Rahmen von MSact! konnte ich erkennen, dass meine pädagogischen Kompetenzen in den letzten Wochen stark angewachsen sind.	3,42	0,67
erf_MSA6_3	Im Rahmen von MSact! konnte ich die Inhalte des Online Campus gut nutzen.	2,75	0,73
erf_MSA7_3	Im Rahmen von MSact! konnte ich lernen, wie man professionelle pädagogische Beziehungen aufbaut.	3,35	0,71
erf_MSA8_3	Im Rahmen von MSact! konnte ich zielorientierte Stundenplanung einüben.	3,28	0,80
erf_MSA9_3	Im Rahmen von MSact! konnte ich lernen, wie ich Schüler/innen für Inhalte begeistere.	3,62	0,58
N		54	

Tabelle 13: Erfahrungen mit dem Trainingsteil MSact! - Deskriptive Statistiken

Die Aussagen, die sich auf praktische Übung beziehen, werden von den teilnehmenden Fellows durchschnittlich mit hoher Zustimmung eingeschätzt (Mittelwerte: 3,28-3,62). Die einzige Ausnahme bildet hier die Anwendung diagnostische Verfahren einzuüben (Mittelwert: 2,58). Im Durchschnitt schätzen sich die teilnehmenden Fellows so ein, dass sie eher nicht mehr zusätzliches pädagogisches Wissen benötigen und empfinden einen starken

Zuwachs ihrer pädagogischen Kompetenzen. Die Fellows beurteilen das erworbene Wissen aus dem Online Campus als eher nutzbar für MSact! (Mittelwert: 2,75).

Die Faktorenanalyse ergibt eine zweifaktorielle Lösung. Dabei laden auf den zweiten Faktor die Items erf\_MSA1, erf\_MSA3 hoch, alle anderen laden auf den ersten Faktor. Allerdings haben Item erf\_MSA6 und erf\_MSA8 auf beiden Faktoren eine fast identische Ladung. Die Reliabilitätsanalyse ergibt auf der Grundlage der bisherigen Skalenbildung keine zufriedenstellenden Werte für Cronbach's alpha. Deshalb wurden die Items hier nicht zu Inhaltsdimensionen zusammengefasst.

#### **7.1.7 Offene Kommentare der Befragten zur Sommerakademie**

Zum dritten Erhebungszeitpunkt (zweite Befragung auf der Katlenburg) wurde zusätzlich zu den standardisierten Fragen mit geschlossenen Antwortoptionen auch eine offene Kommentarmöglichkeit für die Fellows angeboten. Diese war wie folgt eingeleitet: „Zum Abschluss würden wir uns freuen, wenn Sie uns Ihre Ideen und Anmerkungen zur Verbesserung des Trainingsprogramms mitteilen.“ Von dieser Möglichkeit zur freien Äußerung haben 52 der teilnehmenden 62 Fellows Gebrauch gemacht.

Da nach Verbesserungsmöglichkeiten gefragt wurde, überrascht es nicht, dass weit überwiegend aus Teilnehmersicht kritische Aspekte der Sommerakademie thematisiert wurden. Entsprechend darf das Ausbleiben eines Kommentars so gewertet werden, dass sich den betreffenden Fellows zum Befragungszeitpunkt kein Verbesserungsbedarf aufdrängte. Gleichwohl antworten sechs Fellows trotz der Bitte um Kritik ausdrücklich positiv.

Am deutlichsten kommt das Lob für die Organisatoren im folgenden Kommentar zum Ausdruck:

„Insgesamt ist das Trainingsprogramm extrem praxisnah, strukturiert und zielorientiert aufgebaut. Ich habe in den letzten Wochen mehr gelernt als in einigen Semestern an der Uni. Insbesondere die Abstimmung der Inhalte, ständige Reaktivierung sowie die Anwendung in der Praxis haben zu einem enormen Zuwachs der Selbstwirksamkeit bzgl. des pädagogischen Handelns geführt.“ (Fellow 33)

Von den übrigen Fellows wurden weit überwiegend einzelne Phasen der Qualifizierung oder spezifische übergreifende Aspekte der Lernumgebung Sommerakademie kommentiert (Trainer/Trainerinnen; Kommunikation). Weniger Fellows bewerten die Auswahl einzelner Inhalte der Sommerakademie und zwei Fellows äußern sich zu spezifischen Details des Fragebogens. Im Folgenden werden die Verbesserungsideen entsprechend dieser inhaltlichen Schwerpunktsetzung zusammengefasst.

##### **7.1.7.1 Kommentare zu einzelnen Phasen der Qualifizierung**

###### **Online Campus**

- Die inhaltliche Fundierung der Texte im Online Campus wird zum Teil als nicht hinreichend erlebt. Die Ausarbeitung der Inhalte scheint mehrere Fellows Angaben gemäß zu unterfordern, Z.B. „Online Campus konkreter UND theoretischer (Studien



über Erfolgsfaktoren Schule, Lernbehinderungen, Psychologie) anstatt kindergerecht aufbereitete Schlagworte.“

- Auch Forumsaufgaben werden vereinzelt als wenig anregend beschrieben, und das Forum sei als Diskussionsraum ungeeignet (werde auch von TFD nicht als solches forciert). Die geringe Bedeutung der Forumsaufgaben zeige sich auch darin, dass diese nicht aufgegriffen würden.
- Ein individuelles Feedback zu Portfolioaufgaben wird als notwendig für die professionelle Weiterentwicklung erachtet.
- Es wird angeregt, den Online Campus auszubauen, so dass er auch nach der Sommerakademie eine Ressource zu „Schulrecht, Kinderschutz und Schuleinsatz“ biete.
- Weiterhin wird angeregt, die Wichtigkeit von Inhalten mehr hervorzuheben (z.B. notwendiges und erweitertes Wissen zu unterscheiden).

#### *Seminarteil Katlenburg*

- Zum Teil wird die Verschränkung von Online Campus und Akademie kritisiert, in dem Sinne, dass die Akademie auf der Katlenburg als ausschließliche Wiederholung des Online Campus empfunden wurde. Ein Fellow beklagt sich, dass einzelne Inhalte von verschiedenen Trainern / Trainerinnen mehrfach wiederholt worden seien. Der Punkt, dass es auf der Katlenburg zu viel Wiederholung gegeben habe, wird am häufigsten benannt.
- Zugleich wird kommentiert, dass die Inhalte des Online Campus in der Känguru-Woche bzw. in MSAct/Climb nicht genügend aufgegriffen würden.
- Es wird die Idee einer zeitlichen Strukturierung eingebracht, die eine Reflexion von MSAct/Climb im Seminarteil erlaube.
- Insgesamt wird mehr Interaktion und Aktivierung der Fellows (z.B. Rollenspiele, Praxisübungen) und kürzere Frontalphasen gewünscht.

#### *Känguru-Woche*

- Zu diesem Teil der Sommerakademie wird als Verbesserungsmöglichkeit benannt, nicht die Woche vor den Sommerferien dafür auszuwählen.

#### *MSAct/Climb*

- Teilweise scheint es sowohl auf Seiten der Schüler/-innen als auch auf Seiten der Fellows ein Auseinanderklaffen von Erwartungen und tatsächlichem Programm gegeben zu haben. Hier wird mehr Kommunikation als notwendig erachtet.
- Mit Blick auf die Schüler/-innen wird hinterfragt, ob die „zukunftsorientierten“ Teile von allen Schüler/-innen verstanden und gewollt wurden, oder nicht doch mehr fachliche Inhalte gewünscht waren. Vielleicht, so ein anderer Fellow, wäre auch eine stärker schüleraktivierende Gestaltung der Zukunftswerkstatt sinnvoll.



- Mit Blick auf die Teilnehmer/-innen ist zu klären, inwiefern die Anwendung bestehender Lernmaterialien oder auch die Entwicklung eigener Unterrichtssituationen und -materialien eingeübt werden soll.
- Die Unterbringung in Hamburg wird mehrfach, teilweise aus verschiedenen Gründen kritisiert.

#### *7.1.7.2 Weitere Aspekte der Lernumgebung Sommerakademie*

- In den Kommentaren zeigt sich der Wunsch nach einer kontinuierlichen Begleitung durch Trainer, die auch individuelles Feedback geben und mit den Fellows die individuelle Qualifizierung planen.
- Zum Teil wird auch mehr oder stärker begründende Kommunikation gegenüber den Fellows insgesamt gewünscht.
- Es wird eine stärkere Differenzierung für die Primarstufe und die Sekundarstufe gewünscht.

#### *7.1.7.3 Themenbezogene Kommentare*

- Im Fokus der Kritik stehen die angebotenen Inhalte zu Leadership und Kultur, aber auch das ARK-Modell der Sprachförderung. Insbesondere Qualität und Seriosität von Lernangeboten zu Leadership, Kultur und Vision werden angezweifelt; es wird teilweise sogar eine abschreckende Wirkung mit Blick auf die Alumni-Tätigkeit vermutet. Zum Lernangebot Leadership fallen Bemerkungen wie „intellektuelle Beleidigung“, „für Außendarstellung von TFD problematisch“, „wenig Praxisbezug“, „Nähe zum Kitsch“.
- Es werden zusätzliche Angebote gewünscht, erstens zu didaktischen Themen wie Methoden der Leistungsdifferenzierung und Team-Teaching, zweitens zur Fachdidaktik in den Fächern Deutsch, Mathematik und Englisch, sowie drittens zu pädagogischen Themen wie Umgang mit herausfordernden Schüler/-innen, Inklusion, Mobbing, Elternkommunikation. Weiterhin wird ein Angebot zur Sprecherziehung gewünscht. Einzelne Fellows fühlen sich insbesondere in Bezug auf den Umgang mit Schüler/-innen unzureichend vorbereitet.

#### *7.1.7.4 Kommentare zur Befragung*

- Ein Teilnehmer problematisiert die Abfrage von Selbsteinschätzungen, weil sich die Fellows als größtenteils kritisch denkende Menschen kaum „sehr große Fähigkeiten“ attestieren würden. Es wäre sinnvoll, mehr Befragungsformate zu entwickeln, die die Anwendungsfähigkeit von Wissen objektiv erfassen.
- Bei den Fragen zum Belastungserleben werden alternative Antwortmöglichkeiten vorgeschlagen („gering“ bis „sehr hoch“), weil die bestehenden Antwortmöglichkeiten schlecht zur Frage passen. [gute Beobachtung]

Vor dem Hintergrund der zum Teil sehr positiven Bewertungen im standardisierten Teil der Befragung, bieten Kritik und Verbesserungsvorschläge im offenen Antwortformat eine

komplementäre Ressource, um Ansatzpunkte für eine weitere Optimierung der Sommerakademie zu finden.

#### **Vierte Erhebung: Einschätzung der schuleinsatzbegleitenden Maßnahmen**

Der Schuleinsatz begann für die Fellows aufgrund unvorhersehbarer organisatorischer Schwierigkeiten nicht für alle zur gleichen Zeit. Die Hälfte der Fellows, für die hier eine Angabe vorliegt, gibt als Startmonat des Schuleinsatzes den Monat August an. Das heißt für diese Fellows verlief der Start des Einsatzes planmäßig. 13 Fellows geben als Startmonat des Schuleinsatzes den Monat September an, 7 Fellows geben den Monat Oktober an und 4 Fellows geben an, dass ihr Schuleinsatz erst im November begann.

Für die vierte Erhebung liegen je nach Fragenkomplex zwischen 18 und 48 Beobachtungen vor. Durch die geringen Fallzahlen ist oft keine Skalenbildung für die einzelnen Konstrukte möglich.

Zu den TFD Qualifizierungsmaßnahmen während des Schuleinsatzes der Fellows zählen Balkontreffen, Trainer-Fellow-Treffen, Fortbildungen, Hospitation und Bilanzgespräch. Die Balkontreffen werden von den jahrgangsinternen Fellow-Gruppen eigenständig organisiert. Die Fellows haben hier die Gelegenheit, sich im Sinne einer kollegialen Fallberatung gegenseitig auszutauschen und zu unterstützen. Bei Trainer-Fellow-Treffen kommen alle Fellows (auch die Fellows im zweiten Einsatzjahr) eines Trainers/einer Trainerin zusammen<sup>5</sup>, um gegenseitig Informationen auszutauschen und über die eigene Arbeit zu berichten. Bei den Fortbildungen im ersten Einsatzjahr handelt es sich um Weiterbildungen zu pädagogisch-didaktischen Themen. Zur Hospitation kommt der Trainer/die Trainerin des Fellows an dessen Einsatzschule und besucht ihn im Unterricht. Im anschließenden Bilanzgespräch gibt der Trainer/die Trainerin Feedback und bespricht Entwicklungsfelder des Fellows. Dafür wird die TFD Kompetenzmatrix als Diagnoseinstrument zur Gesprächsgrundlage eingesetzt. Diese sollte von den Fellows bezüglich ihrer Qualität und Zweckdienlichkeit eingeschätzt werden. Auf die Frage, wie zufrieden sie mit der Unterstützung von Teach First für ihre Qualifizierung während des Schuleinsatzes seien, zeigen sich die Fellows im Durchschnitt eher zufrieden (Mittelwert: 3,21). Konkrete Verbesserungsvorschläge und Kommentare zur Qualifizierung finden sich unter Abschnitt 7.1.15.

##### **7.1.8 Balkontreffen**

Hinsichtlich der Balkontreffen wurde zunächst erhoben, wie oft die Fellows bereits an diesen teilgenommen haben. Die Antwortmöglichkeiten reichten von 0 Treffen bis 4 Treffen. Insgesamt liegen hier von 48 Fellows Angaben vor. Die meisten Fellows geben an, bereits an zwei Treffen teilgenommen zu haben (39,9%). Ein Fellow hat zum Zeitpunkt seiner Angabe noch an keinem Treffen teilgenommen, 12 Fellows geben an, an einem Treffen, 11 Fellows geben an, an 3 Treffen und 5 Fellows geben an, bereits an 4 Treffen teilgenommen zu haben.

---

<sup>5</sup> Die Zuordnung von Fellows zu Trainern und Trainerinnen erfolgt regional.

Zur Evaluierung der Balkontreffen wurden sechs Items entwickelt. Insgesamt werden die Balkontreffen sehr positiv eingeschätzt. Zu den Einschätzungsfragen der Balkontreffen liegen von 47 Fellows Angaben vor. Am höchsten wird die vertrauensvolle Atmosphäre mit einem Mittelwert von 3,70 bewertet. Die Fellows geben an, dass es eher leicht ist Themen bei den Treffen zu finden, die alle interessieren (Mittelwert: 3,26) und empfinden die Gruppengröße als weder zu groß noch zu klein für den Zweck (Mittelwerte: 1,51; 1,45). Auch den Rückzug der Trainer schätzen die Fellows als eher angemessen ein (Mittelwert: 3,26). Weniger stark zugestimmt wird der Aussage, dass die Teilnehmenden ähnliche Einsatzprofile an den Schulen haben. Dies ist in der Wahrnehmung der Fellows eher nicht der Fall (Mittelwert: 2,21).

Variable	Balkontreffen	MZP (Welle 4)	
		M	STD
balkon1_4	Die TeilnehmerInnen haben ähnliche Einsatzprofile.	2,21	0,93
balkon2_4	Es ist leicht, jeweils Themen zu finden, die alle interessierten.	3,26	0,79
balkon3_4	Die Gruppen sind zu groß für den Zweck.	1,51	0,86
balkon4_4	Die Gruppen sind zu klein für den Zweck.	1,45	0,72
balkon5_4	Der Rückzug der Trainer (etwa nach dem zweiten Treffen) ist angemessen.	3,26	0,85
balkon6_4	Es herrscht eine vertrauensvolle Atmosphäre.	3,70	0,46
N		47	

Tabelle 14: Einschätzung der TFD-Balkon-Treffen - Deskriptive Statistiken

Die Faktorenanalyse ergibt für die Items zur Evaluierung eine dreifaktorielle Lösung. Allerdings laden auf den zweiten und dritten Faktor nur jeweils zwei Items. Auch der Screeplot spricht für eine einfaktorielle Lösung. Eine Reliabilitätsanalyse der Items des ersten Faktors (balkon1, balkon2, balkon5, balkon6) zeigt allerdings keine zufriedenstellende Reliabilität (Cronbach's alpha: 0,40). Dies liegt an einer sehr geringen Trennschärfe aller Items. Dadurch ist hier keine Skalenbildung möglich.

#### 7.1.9 Trainer-Fellow-Treffen

Wie auch bei den Balkontreffen sollten die Fellows zunächst angeben wie oft sie bereits an einem Trainer-Fellow-Treffen (TuFT) teilgenommen haben. Die Antwortmöglichkeiten reichten erneut von 0 Treffen bis 4 Treffen. Insgesamt haben 48 Fellows hier eine Angabe gemacht. Die meisten Fellows geben an, bereits an drei Treffen teilgenommen zu haben (33,3%). Vier Fellows haben zum Zeitpunkt ihrer Angabe noch an keinem Treffen teilgenommen, 9 Fellows geben an, an einem Treffen, 8 Fellows geben an, an zwei Treffen und 11 Fellows geben an, bereits an 4 Treffen teilgenommen zu haben.

Zur inhaltlichen Einschätzung der Trainer-Fellow-Treffen wurden vier Items entwickelt. Insgesamt haben 44 Fellows zu diesem Fragenkomplex Angaben gemacht. Im Durchschnitt werden die Items weder besonders positiv noch besonders negativ eingeschätzt. Die Mittelwerte liegen alle um den theoretischen Mittelwert von 2,50 herum. Das heißt die Antworten liegen im Durchschnitt genau zwischen „trifft eher nicht zu“ und „trifft eher zu“.



Die Fellows geben an, dass die Treffen eher einen hohen Informationsgehalt haben (Mittelwert: 2,75), aber sie geben im Durchschnitt auch an, dass weder unterrichtsrelevante Themen, noch bildungsrelevante Themen, noch Fellow-Schüler-Interaktion besonders intensiv besprochen wurden (Mittelwerte: 2,32-2,52).

Variable	Trainer-Fellow-Treffen	MZP (Welle 4)	
		M	STD
trainertreff1_4	Die Treffen haben einen hohen Informationsgehalt.	2,75	0,78
trainertreff2_4	Unterrichtsrelevante Themen werden intensiv besprochen.	2,32	0,74
trainertreff3_4	Bildungsrelevante Themen werden intensiv besprochen.	2,52	0,66
trainertreff4_4	Fellow-Schüler-Interaktion wird intensiv besprochen.	2,41	0,69
N		44	

Tabelle 15: Einschätzung der TFD-Trainer-Fellow-Treffen - Deskriptive Statistiken

Die Faktorenanalyse ergibt für die Items zu den Trainer-Fellow-Treffen eine zweifaktorielle Lösung, wobei auf den zweiten Faktor nur das Item trainertreff3 hoch lädt. Auch der Screeplot spricht für eine einfaktorielle Lösung. Die anschließende Reliabilitätsanalyse zeigt eine nicht mehr zufriedenstellende Reliabilität für die drei Items des ersten Faktors (Cronbach's alpha: 0,72). Dadurch ist hier keine Skalenbildung möglich.

#### 7.1.10 Fortbildungen

Den Fragen zur Einschätzung der TFD Fortbildungen ging wieder eine Frage zur Häufigkeit der Teilnahme voraus. Die Antwortmöglichkeiten waren 0, 1 und 2. Von den 48 Fellows, die hier eine Angabe machen, haben die meisten zum Zeitpunkt ihrer Angabe bereits an zwei Fortbildungen teilgenommen (72,9%), 12 Fellows geben an, an einer Fortbildung, und ein Fellow gibt an, noch an keiner Fortbildung teilgenommen zu haben.

Zur Evaluierung der TFD Fortbildungen wurden insgesamt 10 Items entwickelt. Diese fokussieren auf die inhaltliche und methodische Gestaltung der Fortbildungen. Eingeleitet wurden die Fragen mit „Während der bisher besuchten Fortbildungen...“. Zu den Fortbildungsfragen haben 47 Fellows Angaben gemacht. Insgesamt sind die Einschätzungen der Fortbildungen sehr positiv. Sechs von zehn Items werden im Durchschnitt deutlich über 3,0 bewertet. Am stärksten wird der Aussage zugestimmt, dass die Fellows sich gegenseitig beim Lernen unterstützen (Mittelwert: 3,51). Auch die Aussagen, dass durch Fragen aus der Gruppe neue Lösungsansätze für Schwierigkeiten in Unterrichtssituationen entstanden sind und dass das Wissen aus der Sommerakademie vertieft wurde, wurden im Durchschnitt sehr hoch eingeschätzt (Mittelwerte: 3,43; 3,30). Weiter schätzen sie die Aussagen, dass sie mit anderen Fellows gemeinsam Fragen pädagogischen Handelns diskutiert, sich über pädagogische Handlungsalternativen ausgetauscht und über Lösungen in schwierigen Unterrichtssituationen beratschlagt haben, sehr positiv ein (Mittelwerte: 3,40-3,45). Die Aussagen, dass das Gelernte an praktischen Unterrichtsbeispielen geübt werden konnte, dass Fragen besprochen wurden, die sich aus konkreten Situationen während des Schuleinsatzes ergaben und dass Handlungssicherheit für Unterrichtssituationen durch



konkrete Übungen gewonnen werden konnte, werden im Mittel mit *trifft eher zu* eingeschätzt (Mittelwerte: 2,87-3,04). Am wenigsten zutreffend wird die Aussage eingeschätzt, dass sich die Fellows gegenseitig Feedback zur Entwicklung ihrer pädagogischen Kompetenzen gegeben haben (Mittelwert: 2,70).

Variable	Fortbildungen	MZP (Welle 4)	
		M	STD
	<b>Einleitungstext:</b> Während der bisher besuchten Fortbildungen...		
fortbildung1_4	... haben sich die Fellows gegenseitig beim Lernen unterstützt.	3,51	0,62
fortbildung2_4	... entstanden durch Fragen aus der Gruppe neue Lösungsansätze für Schwierigkeiten in Unterrichtssituationen.	3,43	0,74
fortbildung3_4	... konnte ich das bisher Gelernte an praktischen Unterrichtsbeispielen üben.	2,89	0,91
fortbildung4_4	... haben sich die Fellows gegenseitig Feedback zur Entwicklung ihrer pädagogischen Kompetenzen gegeben.	2,70	0,95
fortbildung5_4	... wurde das Wissen aus der Sommerakademie vertieft.	3,30	0,75
fortbildung6_4	... wurden meine Fragen, die sich aus konkreten Situationen während des Schuleinsatzes ergaben, besprochen.	2,87	0,80
fortbildung7_4	... konnte ich durch konkrete Überlegungen Handlungssicherheit für Unterrichtssituationen gewinnen.	3,04	0,83
fortbildung8_4	... habe ich mit anderen Fellows gemeinsam Fragen pädagogischen Handelns diskutiert.	3,40	0,68
fortbildung9_4	... habe ich mich mit anderen Fellows über pädagogische Handlungsalternativen ausgetauscht.	3,45	0,65
fortbildung10_4	... habe ich mit anderen Fellows über Lösungen in schwierigen Unterrichtssituationen beratschlagt.	3,40	0,68
N		47	

Tabelle 16: Einschätzung der TFD-Fortbildungen - Deskriptive Statistiken

Die Faktorenanalyse ergibt eine dreifaktorielle Lösung, wobei auf den zweiten Faktor nur Items laden, die eine praktisch identische oder höhere Ladung auf dem ersten Faktor aufweisen. Für den dritten Faktor trifft dies auf eines von zwei Items zu, sodass hier nur ein Item mit hoher Ladung lädt. Der Screeplot spricht ebenfalls für eine einfaktorielle Lösung. Auf den ersten Faktor laden die Items fortbildung1 – fortbildung3 und fortbildung5 – fortbildung10. Die anschließende Reliabilitätsanalyse zeigt eine gute Reliabilität an (Cronbach's alpha: 0,86). Die neu gebildete Skala *Fortbildung* hat einen Mittelwert von 3,26.

#### 7.1.11 Hospitation/Bilanzgespräch

Zu den Lerngelegenheiten Hospitation und Bilanzgespräch wurde lediglich eine Frage nach der Häufigkeit der Teilnahme gestellt, da zum Start der vierten Befragung einige Fellows noch nicht lange an einer Schule im Einsatz waren. Die meisten Fellows (85,1%) wurden bis zum Zeitpunkt der Angabe ein (48,9%) oder zweimal (36,2%) von ihrem Trainer oder ihrer Trainerin im Unterricht besucht (Hospitation). Drei Fellows (6,4%) gaben an, dass sie noch keinen Unterrichtsbesuch hatten.

Die Mehrheit der Fellows hat zum Zeitpunkt der Angabe noch nicht an einem Bilanzgespräch teilgenommen (61,7%). 17 Fellows (36,2%) haben bereits an einem Bilanzgespräch und ein Fellow (2,1%) an zwei Bilanzgesprächen teilgenommen.

An die Frage zum Bilanzgespräch wurde die Frage angeschlossen, ob die TrainerInnen im Gespräch die Kompetenzmatrix nutzten. Diese Frage war als Filterfrage programmiert, d.h. alle Fellows, die hier eine andere Antwort als „nie“ eingaben, wurden zu den Fragen zur Kompetenzmatrix weitergeleitet. Für die Fellows, die angaben, dass die TrainerInnen noch nie die Kompetenzmatrix genutzt hätten, wurde der Fragenblock zur Kompetenzmatrix übersprungen. Dabei handelt es sich um den Großteil der Fellows. Dies ergibt sich aus der Tatsache, dass nur 18 Fellows an Bilanzgesprächen teilgenommen haben. Die Evaluation der Kompetenzmatrix beruht demnach nur auf diesen 18 Personen.

#### 7.1.12 Kompetenzmatrix

Zur Evaluation des Diagnoseinstruments Kompetenzmatrix wurden 6 Items entwickelt. Davon sind zwei negativ formuliert (matrix2\_4 und matrix4\_4), sodass hier niedrige Werte positiv zu interpretieren sind.

Variable	Kompetenzmatrix	MZP (Welle 4)	
		M	STD
matrix_1_4	Die Kompetenzmatrix erweist sich als gutes Instrument zur Gesprächsvorbereitung.	2,94	0,80
matrix_2_4	Die Matrix erweist sich als wenig nützlich, um die weitere Kompetenzentwicklung zu planen.	1,78	0,73
matrix_3_4	Mir ist klar, welches Handlungswissen von mir auf den einzelnen Kompetenzstufen erwartet wird.	2,78	0,73
matrix_4_4	Es gibt wichtige Entwicklungsbereiche, die von der Kompetenzmatrix nicht abgedeckt werden.	2,17	0,79
matrix_5_4	Die Matrix erweist sich als gutes Instrument um individuelle Veränderungen festzuhalten.	2,89	0,76
matrix_6_4	Die einzelnen Kompetenzstufen sind gut voneinander zu unterscheiden.	2,44	0,78
N		18	

Tabelle 17: Einschätzung der TFD-Kompetenzmatrix - Deskriptive Statistiken

Im Durchschnitt bewerten die Fellows die Kompetenzmatrix eher gut – die Mittelwerte der meisten Items liegen um 3,0 herum<sup>6</sup>. So geben die Fellows an, dass die Matrix sich eher als gutes Instrument zur Gesprächsvorbereitung und zum Festhalten der individuellen Veränderung eignet (Mittelwerte: 2,94; 2,89). Die Fellows lehnen die Aussagen eher ab, dass die Matrix sich als wenig nützlich zur Planung der Kompetenzentwicklung erweist und dass es wichtige Entwicklungsbereiche gibt, die von der Kompetenzmatrix nicht abgedeckt

<sup>6</sup> Die Werte der negativ gepolten Items müssen zur Berechnung der Reliabilität sowie der neuen Skala umgerechnet werden. Dazu werden alle Werte vom Wert 5 abgezogen (Summe aus kleinst- und größtmöglichem anzukreuzenden Wert). Dadurch ergibt sich als Mittelwert für matrix2\_4 ein Wert von 3,29 und für matrix4\_4 ein Wert von 2,86.

werden (Mittelwerte: 1,78; 2,17). Den Fellows ist auch eher klar, welches Handlungswissen auf welcher Kompetenzstufe von ihnen erwartet wird (Mittelwert: 2,78). Das Item, das am schlechtesten eingeschätzt wird, betrifft die Konzeption des Instrumentes selbst. So können die Fellows die einzelnen Kompetenzstufen eher nicht gut voneinander unterscheiden (Mittelwert: 2,44).

Bei der Faktorenanalyse ergibt sich eine zweifaktorielle Lösung, wobei auf den zweiten Faktor nur die beiden negativ formulierten Items laden. Davon hat matrix\_4 die deutlich höhere Ladung auf dem ersten Faktor, sodass auf den zweiten Faktor nur noch matrix\_4 lädt. Auch der Screeplot spricht für eine einfaktorielle Lösung. Die Überprüfung der Reliabilität der Items des ersten Faktors ergibt ein Cronbach's alpha von 0,74, was durch Ausschluss von Item matrix\_6, welches eine schlechte Trennschärfe aufweist, auf eine akzeptable Reliabilität von 0,76 verbessert werden kann. Das heißt in die neue Variable fließen vier der sechs ursprünglichen Items ein. Der Mittelwert der neu gebildeten Variable *Matrix* beträgt 2,60.

### 7.1.13 TFD Werkstatt

Die TFD Werkstatt ist eine Online Plattform auf der TFD intern Informationen und Materialien ausgetauscht werden können. Es wurden 7 Items entwickelt, um die *Übersichtlichkeit* und technische *Funktionalität* der Plattform zu evaluieren und eine offene Frage gestellt, wie die Werkstatt noch verbessert werden könnte. Die Auswertung der Fragen, die zur *Nutzung der Werkstatt* gestellt wurden, finden sich bei den Variablen zur Nutzung der Lernumgebung unter Abschnitt 7.2.4.

Variable	TFD Werkstatt (Übersichtlichkeit/technische Funktionalität)	MZP (Welle 4)	
		M	STD
werkstatt1_4	Die Navigation auf den Seiten der Werkstatt funktioniert gut.	2,27	0,71
werkstatt2_4	Die Nutzung der Werkstatt ist selbsterklärend.	2,38	0,70
werkstatt3_4	Die Ordnerstruktur der Werkstatt ist übersichtlich gestaltet.	1,98	0,70
werkstatt4_4	Die Forenstruktur der Werkstatt ist übersichtlich gestaltet.	2,06	0,78
werkstatt5_4	Die Werkstatt-Seiten sind überladen mit Informationen.	2,71	0,71
werkstatt6_4	Es ist technisch einfach, einen Beitrag im Forum zu schreiben.	3,25	0,67
werkstatt7_4	Es ist technisch einfach, ein Dokument hochzuladen.	3,21	0,77
N		48	

Tabelle 18: Nutzerfreundlichkeit der TFD-Werkstatt

Zum TFD Werkstatt Fragenkomplex haben 48 Fellows Angaben gemacht. Die Items, welche die *Übersichtlichkeit* der Werkstatt betreffen, werden von den Fellows eher negativ bewertet. So nehmen die Fellows die Navigation auf den Seiten der Werkstatt als eher nicht gut funktionierend wahr und empfinden auch die Nutzung der Werkstatt als eher nicht selbsterklärend (Mittelwerte: 2,27; 2,38). Sowohl die Ordner als auch die Forenstruktur wird von den Fellows als eher nicht übersichtlich gestaltet eingeschätzt (Mittelwerte: 1,98; 2,06). Im Gegensatz dazu werden die Werkstatt-Seiten als eher nicht informationsüberladen



eingeschätzt (Mittelwert: 2,71) und auch die *technische Funktionalität* (Beitrag im Forum schreiben; Dokument hochladen) wird eher gut bewertet (Mittelwerte: 3,25; 3,21).

Die Faktorenanalyse ergibt eine dreifaktorielle Lösung, wobei auf den zweiten Faktor nur Item werkstatt5 und auf den dritten Faktor Item werkstatt6 und werkstatt7 hoch laden. Auch der Screeplot spricht für nur einen Faktor. Die Reliabilitätsanalyse der Items des ersten Faktors zeigt eine noch zufriedenstellende Reliabilität (Cronbach's alpha: 0,76), die aber unter Ausschluss von Item werkstatt2, welches eine schlechte Trennschärfe aufweist, deutlich verbessert werden kann (Cronbach's alpha: 0,80). Die neue Variable *Übersichtlichkeit\_Werkstatt*, gebildet aus den Items werkstatt1, werkstatt3 und werkstatt4 hat einen Mittelwert von 2,10.

### *Verbesserungsvorschläge für die TFD-Werkstatt*

Zusätzlich zu den Einschätzungsfragen zur TFD Werkstatt gab es die Möglichkeit, zur Frage „Wie könnte man die Werkstatt noch besser gestalten?“ Verbesserungsvorschläge und Kommentare abzugeben. Insgesamt haben hier 22 von 48 Fellows Angaben gemacht. Die meisten Fellows, die zu dieser Frage Angaben machen, nennen mehrere Aspekte, die aus ihrer Sicht optimierbar wären. Dadurch gibt es insgesamt mehr als 22 Verbesserungsvorschläge bzw. Kommentare. Dabei betrifft der Großteil der Vorschläge die Gestaltung der Werkstatt (Übersichtlichkeit, Funktionen), aber es gibt auch Verbesserungsvorschläge zu den Inhalten der Seite (Aktualisierung von Inhalten, Bewertung von Inhalten, Wunschinhalte) und zur Nutzung der Werkstatt (z.B. Zugriffsrechte, Nutzung steigern). Im Folgenden werden die Verbesserungsvorschläge und Kommentare zu den oben genannten thematischen Schwerpunkten zusammengefasst.

### *Gestaltung der Werkstatt*

- Zu diesem Thema werden Vorschläge gemacht, wie die Werkstatt übersichtlicher und strukturierter gestaltet werden könnte. Hierzu gibt es insgesamt 10 Kommentare (z.B.: „Zu allererst braucht es ein hochwertiges User Design, das eine einfache Navigation und eine gute Übersicht über die Angebote auf der Werkstatt bietet.“; „Noch übersichtlicher machen: z.B. Unterteilung nach - Schulform, Klassenstufe bzw. Leistungsstufe, Fach, Thema.“; „Die aktuellen Beiträge im Forum sollten chronologisch abwärts angeordnet sein. Das Aktuellste sollte somit ganz oben, zuerst, stehen!!!!!!“; „Die Werkstatt könnte zu einer Dropbox ausgebaut werden.“; „Es wäre ausgesprochen hilfreich, wenn die Dokumente und Ordner Titel hätten, aus denen man bereits schließen kann, worum es geht.“; „Die Trennung in Ordner und Foren empfinde ich oft als verwirrend.“)
- Des Weiteren werden Wünsche für zusätzliche Funktionen geäußert. Relativ häufig wird der Wunsch nach verschiedenen Suchfunktionen geäußert (z.B.: „Suchfunktion nach Stichwörtern (Wiki)“; „Eine großartige Vorlage bietet "meinunterricht.de". Hier ermöglicht eine tiefe Verschlagwortung von Beiträgen eine gezielte Suche.“; „Fellow-Suchfunktion mit Funktionen: Diese liefert dann z.B. alle Fellows zurück, die "in



NRW", "in Klasse 6", "DAZ" unterrichtet haben, so dass man diese dann gezielt anschreiben kann.“).

- Zusätzlich werden verschiedene andere Funktionen gewünscht (z.B. „Kalenderfunktion nutzen“; „Wikieinrichtung“; „Einrichtung eines Bewertungssystems“; „Antwortfunktion zu einem bestimmten Beitrag; Benachrichtigung, wenn jemand auf den eigenen Kommentar antwortet.“; „Anzeige im Profil, welche Dateien der Fellow hochgeladen hat.“; „Ein Cloud-Speicher, damit man Versionen verwalten kann und besser zusammen Arbeitsblätter erstellen kann.“; „Schön wäre auch, wenn man alle Dokumente in einer Vorschau sehen könnte ohne sie direkt runterzuladen.“; „Beiträge dieses Nutzers"- und "Dateien dieses Nutzers"-Button.“; „Seitennavigation!!! Also eine Randspalte, die immer sichtbar ist und verdeutlicht, wo man sich befindet.“; „Benachrichtigungen differenzierter gestalten, man bekommt entweder alles oder keine Nachrichten.“).

#### Inhalte der Werkstatt

- Es werden Vorschläge gemacht, die die inhaltliche Wartung betreffen (z.B. „Administratoren könnten Unterrichtsmaterial besser sortieren, auch nach Alter (Grundschule oder weiterführende Schule).“; „Auswertung der Informationen durch Mitfellows“; „Zuständiges Kleinteam für die Sammlung und Verwaltung von Lernmaterialien. Kein eigenes Hochladen.“).
- Außerdem werden Vorschläge dazu gemacht, welche Inhalte auf der Werkstatt zur Verfügung gestellt werden (z.B. „mehr Integration der Lehrmaterialien von FuF“; „Links, die aktuell relevant sind (z.B. Wirkungsmessungsmatrizen) prominent platzieren.“; „Bereitstellung von mehr Arbeitsmaterialien bzw. Vorstellung von Links und Arbeitsmaterialien; Bücher- und Materialbasar; ... “).

#### Nutzung der Werkstatt

- Es wird von einigen angemerkt, dass die Werkstatt mehr genutzt werden sollte (z.B. „Eigentlich schade, dass alle ständig Material anfertigen und man so wenig gegenseitig davon profitiert.“).
- Es wird angemerkt, dass es umständlich ist auf bestimmte Ordner oder Gruppen zuzugreifen (z.B. „Triviale Zugriffsrechte auf Gruppen nicht erst auf Anfrage sondern von Beginn an.“; „Es dauert noch lange bis man freigeschaltet wird für eine Gruppe.“).
- Oder, dass der Zugriff Passwort geschützt sein sollte („Verschlüsselungsmöglichkeiten / Passworteinrichtung bei Material, dass nur für bestimmte Gruppen bestimmt ist (Fotos, Material), allgemein Datenschutz (Schulmaterial/SuS!)“).

#### 7.1.14 Lernumgebung Schule

Zur Evaluation der Lernumgebung „Einsatzschule“ wurden Items zur sozialen Eingebundenheit der Fellows in der Schule entwickelt. Weiterhin wurde das Einsatzprofil der Fellows erhoben. Die Fellows sollten hier angeben, wie sich ihre Zeit an der Einsatzschule auf verschiedene Tätigkeitsbereiche verteilt.

### 7.1.14.1 Soziale Eingebundenheit

Zunächst sollten die Fellows angeben, ob es an ihrer Einsatzschule eine oder mehrere Personen gibt, durch die sie betreut werden – ähnlich der Person eines Mentors, oder ob sie keine solche Betreuung erfahren. Insgesamt geben 17 Personen (35,4%) an, dass sie von einer Person betreut werden, 14 (29,2%) Personen geben an, dass sie von mehreren Personen betreut werden und 17 Personen (35,4%) geben an, dass sie keine Betreuungsperson haben. Zusätzlich wurden sechs Items zur sozialen Eingebundenheit an der Schule entwickelt. Diese werden von den Fellows, die angeben, dass sie eine oder mehrere Betreuungspersonen haben, wie zu erwarten positiver eingeschätzt als von den Fellows, die angeben, dass sie keine Betreuungsperson haben. Tabelle 19 zeigt die Einschätzungen getrennt für die Fellows, mit und ohne Betreuungsperson.

Variable	Soziale Eingebundenheit an Einsatzschule	MZP (Welle 4)			
		Fellows mit Betreuungsperson		Fellows ohne Betreuungsperson	
		M	STD	M	STD
	<b>Einleitungstext:</b> Es gibt eine oder mehrere Lehrpersonen,...				
sozeinbind1_4	... die mir konkret Hilfestellung geben, wenn ich nicht weiß, wie ich mich in schwierigen Situationen verhalten soll.	3,32	0,83	2,94	1,03
sozeinbind2_4	... die mit mir Unterrichtsmaterialien austauschen.	3,23	0,99	2,82	1,01
sozeinbind3_4	... die ich ansprechen kann, wenn ich Feedback zu meinen Unterrichtsmaterialien einholen möchte.	3,35	0,88	2,76	1,15
sozeinbind4_4	... mit denen ich mich gewinnbringend über den Umgang mit problematischen Schülern und Schülerinnen austauschen kann.	3,48	0,77	3,00	1,12
sozeinbind5_4	... mit denen ich mich über Unterrichtsstrategien/-Methoden austauschen kann.	3,10	0,87	3,00	1,00
sozeinbind6_4	... die mir aktiv helfen, neue Ideen/Projekte umzusetzen.	2,74	0,86	2,24	1,15
N		31		17	

Tabelle 19: Einschätzung der sozialen Eingebundenheit mit versus ohne Betreuungsperson - Deskriptive Statistiken

Die Gruppe der Fellows mit Betreuungsperson(en) schätzt ihre soziale Eingebundenheit recht hoch ein. Die Mittelwerte liegen für fünf von sechs Items deutlich über 3,0. So geben die Fellows im Durchschnitt an, dass es eine oder mehrere Lehrpersonen gibt, die ihnen Hilfestellung geben, wenn sie in schwierigen Situationen nicht sicher sind, wie sie sich verhalten sollen, mit denen sie Unterrichtsmaterialien austauschen und die sie ansprechen können, wenn sie Feedback zu ihren eigenen Unterrichtsmaterialien einholen möchten

(Mittelwerte: 3,32; 3,23; 3,35). Am stärksten stimmen die Fellows der Aussage zu, dass sie sich mit anderen Lehrpersonen gewinnbringend über den Umgang mit problematischen Schülern und Schülerinnen austauschen (Mittelwert: 3,48), etwas niedriger schätzen sie den Austausch mit anderen Lehrpersonen über Unterrichtsstrategien/-Methoden ein (Mittelwert: 3,10). Am wenigsten stimmen sie der Aussage zu, dass es eine oder mehrere Lehrpersonen gibt, die ihnen aktiv helfen, neue Ideen/Projekte umzusetzen (Mittelwert: 2,74).

Schaut man sich die Gruppe der Fellows an, die angeben, dass sie keine Betreuungsperson(en) an der Einsatzschule haben, dann fällt auf, dass die Mittelwerte zwar niedriger ausfallen als bei der Gruppe mit Betreuungsperson(en), aber trotzdem fast alle um 3,0 herum liegen. Das heißt, dass im Durchschnitt auch die Fellows dieser Gruppe die Aussagen zur Einschätzung der sozialen Eingebundenheit als *eher zutreffend* eingeschätzt haben. Die Aussagen mit der größten und der kleinsten Zustimmung sind in beiden Gruppen die gleichen.

Die Faktorenanalyse ergibt eine einfaktorielle Lösung. Sämtliche Items haben hohe Ladungen auf dem ersten Faktor. Die niedrigste Ladung hat mit 0,70 das Item sozeinbind6. Die anschließende Reliabilitätsanalyse ergibt ein sehr gute Reliabilität (Cronbach's alpha: 0,91). Der Mittelwert der neugebildeten Skala *soziale Einbindung* beträgt 3,06, d.h. im Durchschnitt sind die Fellows an der Einsatzschule eher sozial eingebunden.

#### **7.1.14.2 Einsatzprofil/Tätigkeiten der Fellows**

Die Einsatzprofile der Fellows und damit auch ihre Lerngelegenheiten können je nach Einsatzschule unterschiedlich gestaltet sein. Um einen Überblick über die verschiedenen Tätigkeiten zu erhalten, wurde den Fellows eine Liste mit 12 Tätigkeiten vorgegeben. Zu jeder Tätigkeit sollten sie angeben, wie viele Arbeitsstunden sie in einer normalen Schulwoche auf diese Tätigkeit verwenden. Zusätzlich zu den 12 vorgegebenen Tätigkeiten gab es die Möglichkeit unter „Sonstiges“ weitere Tätigkeiten anzugeben, die nicht durch die Liste abgedeckt sind. Tabelle 20 zeigt die durchschnittlichen Stunden pro Schulwoche, die ein Fellow auf die verschiedenen Tätigkeiten verwendet. Um die Spannweite der Stundenverteilung zu verdeutlichen, sind in der Tabelle ebenfalls der jeweils kleinste und der größte Stundenwert (Min; Max) angegeben.

Interessanterweise gibt es, bis auf die Tätigkeit *Gespräche mit anderen Lehrpersonen führen* (Mittelwert: 2,54), keine Tätigkeit, die von allen Fellows ausgeübt wird. Die beiden Tätigkeiten, die von den Fellows mit dem mit Abstand größten Stundenumfang ausgeübt werden, sind *Unterrichten einer Teilungsgruppe als alleinige Lehrperson* und *Teamteaching* (Mittelwerte: 6,90; 9,17). Bei diesen Tätigkeiten ist auch die Stundenspanne am größten: es gibt Fellows, die diese Tätigkeiten gar nicht ausüben, während andere Fellows diesen 22 Stunden pro Schulwoche nachgehen. Im Durchschnitt etwa zwei bis zweieinhalb Stunden pro Schulwoche verwenden die Fellows auf die Tätigkeiten *Gespräche mit Lehrpersonen* (Mittelwert: s.o.) bzw. *mit Schülern/Schülerinnen außerhalb des Unterrichts führen*



(Mittelwert: 2,02), *Unterricht für eine ganze Klasse halten* (Mittelwert: 1,90), *Spezifische Förderangebote außerhalb des Unterrichts* (Mittelwert: 1,98) und *Projekte* (Mittelwert: 2,42). Die übrigen Tätigkeiten werden im Durchschnitt weniger als eine Stunde ausgeübt.

Variable	Lerngelegenheiten/Einsatzprofil der Fellows	MZP (Welle 4)			
		M	STD	Min	Max
	<b>Einleitungstext:</b> In einer normalen Schulwoche: Wie viele Arbeitsstunden verwenden Sie jeweils auf die folgenden Tätigkeiten:				
einsatzzeit1_4	Gespräche mit einzelnen Schülern/Schülerinnen oder Schülerkleingruppen außerhalb des Unterrichts	2,02	1,30	0,00	5,00
einsatzzeit2_4	Gespräche mit Lehrkräften oder weiteren Professionellen	2,54	1,20	1,00	5,00
einsatzzeit3_4	Gespräche mit Eltern	0,33	0,56	0,00	2,00
einsatzzeit4_4	Unterricht einer Teilungsgruppe als alleinige Lehrperson	6,90	5,40	0,00	22,00
einsatzzeit5_4	Teamteaching	9,17	5,23	0,00	22,00
einsatzzeit6_4	Unterricht für eine ganze Klasse (gleichsam als Lehrkraft)	1,90	2,50	0,00	10,00
einsatzzeit7_4	Spezifische Förderangebote außerhalb des Unterrichts mit relativ verbindlicher Teilnahme	1,98	2,05	0,00	7,00
einsatzzeit8_4	Offene Angebote zur Lernberatung	0,31	0,72	0,00	3,00
einsatzzeit9_4	Weitere offene Angebote	0,88	1,21	0,00	5,00
einsatzzeit10_4	Berufsberatung	0,33	1,00	0,00	5,00
einsatzzeit11_4	Projekte	2,42	2,24	0,00	10,00
einsatzzeit12_4	Schulentwicklung	0,92	1,11	0,00	5,00
einsatzzeit13_4	Sonstiges	1,77	3,32	0,00	17,00
N		48			

Tabelle 20: Lerngelegenheiten/Einsatzprofil der Fellows - Deskriptive Statistiken

Unter der Kategorie Sonstiges haben insgesamt 20 Fellows eine bzw. mehrere zusätzliche Angaben gemacht. Darunter fallen die Vorbereitung von Unterricht (z.B. Beschaffung und Vorbereitung von Unterrichtsmaterialien, Lernzeit), Betreuungsaufgaben von Schülern (z.B. Betreuung verspäteter Schüler, Einzelcoaching von Schülern im Trainingsraum, die den Unterricht gestört haben, Lernbüro, Einzelförderung parallel zum Unterricht, Teilnahme an Klassenfahrten, Bewegungspause, Pausenaufsichten), Organisatorische Aufgaben (z.B. Planung des Schuljubiläums, Fundraising) und Besprechungen (z.B. Treffen mit der Schulsozialarbeiterin). Für diese Tätigkeiten verwenden die Fellows im Durchschnitt etwa 2 Stunden pro Schulwoche (Mittelwert: 1,77). Wären diese Tätigkeiten in die Liste der Tätigkeiten aufgenommen worden, so hätte sich aller Wahrscheinlichkeit nach ein durchschnittlich höherer Gesamtstundensatz ergeben.

#### 7.1.15 Offene Kommentare zur Qualifizierung während des Schuleinsatzes

In der vierten Erhebung gab es für die Fellows die Möglichkeit, Verbesserungsvorschläge zur Optimierung der Qualifizierung abzugeben. Die offene Frage im Online Instrument lautete:



„Wie würden Sie die Qualifizierung während des Schuleinsatzes durch TFD verbessern wollen?“

Zu dieser Frage haben insgesamt 34 Personen Verbesserungsvorschläge bzw. Kommentare abgegeben. Bei den Fellows, die zu dieser Frage keine Angaben machen, ist davon auszugehen, dass sie keinen Verbesserungsbedarf des Programms sehen. Ein Fellow spricht sogar sein ausdrückliches Lob aus: „Ich bin mit der Qualifizierung sehr zufrieden! Ich finde, dies ist eins der Dinge, die TFD am besten macht!“.

Die übrigen Kommentare beziehen sich auf Aspekte des Qualifizierungsprogramms insgesamt (Kommentare zu Didaktik, Methoden und Inhalten), auf einzelne schuleinsatzbegleitende Maßnahmen (TuFTs, Fortbildungen, Hospitationen und Bilanzgespräche), den Schuleinsatz (Einführung der Fellows an den Schulen, Kooperation mit Schulleitung), Professionelle Kooperation (zwischen Fellows, zwischen Lehrenden und den Schulen und Fellows), Möglichkeiten der Zertifizierung der Fellows sowie Belastungsfaktoren (Arbeitsbelastung, Erwartungsdruck). Es folgt eine Zusammenfassung der Kommentare zu den oben genannten inhaltlichen Schwerpunkten.

#### **7.1.15.1 Kommentare zum Qualifizierungsprogramm insgesamt**

- Insgesamt wünschen sich viele Fellows ein Mehr an Fachliteratur und Didaktik, das teilweise auch spezifisch ausformuliert wird (z.B. „Im Vorhinein mehr und oder anders strukturierte Vermittlung (inkl. Übungsphasen) der Kompetenzen zur Unterrichtsplanung, Strukturierung, Differenzierung und Klassenmanagement“; „Es wäre sinnvoll, in der Ausbildung noch mehr praktisch zu unterrichten, damit der Schritt von langfristiger zu kurzfristig Planung zur Umsetzung ausreichend geübt wurde, wenn man anfängt in der Schule zu unterrichten.“).
- Einige Fellows äußern Themen, die sie für besonders wichtig halten (z.B. „Elternarbeit thematisieren“; „Texte über Inklusion oder Texte, wie man mit schwierigen Schülern umgeht“)
- Zwei Fellows äußern sich zu den Trainern und Trainerinnen („Ich habe immer noch das Gefühl, dass wenig pädagogische Expertise bei TFD herrscht. Schaltet doch mal den Herrn [...] ein. Die Einheit zur Differenzierung empfand ich jedenfalls als sehr gewinnbringend und frage mich, warum wir nicht schon früher über verschiedene Differenzierungsmöglichkeiten geredet haben“; „Teilweise habe ich das Gefühl, dass die TrainerInnen am Rande der Überlastung stehen und sich das auf die Qualifizierung auswirkt.“).
- Von einigen Fellows wird eine Schärfung des Einsatzprofils vermisst (z.B. „Meiner Meinung nach werden zu viele Dinge auf einmal verlangt. Das Profil, was Fellows machen sollen finde ich durch TFD noch nicht geschärft.“; „Eine noch stärkere Fokussierung meines Arbeitsprofils auf Lerngruppen würde ermöglichen, dass ich mich stärker qualifiziere, indem ich nicht als "Wachmann"/"Feuerwehrmann" im Klassenzimmer eingesetzt werde, sondern stärker in die Planung und gezielte Diagnostik einsteigen könnte.“).

### 7.1.15.2 Kommentare zu einzelnen schuleinsatzbegleitenden Maßnahmen

#### TuFTs

- Es werden allgemein formulierte Kommentare abgegeben, die das Gesamtkonzept der TuFTs bemängeln (z.B. „Die Trainer und Fellow Treffen finde ich konzeptuell noch nicht ausgereift“; „Die TuFTs sollten überarbeitet werden.“; „Mit den TuFTs bin ich noch nicht ganz zufrieden, da oft Zeit verloren geht und der Mehrwert für mich gering ist bei dieser Veranstaltung.“).
- Des Weiteren äußern die Fellows Wünsche nach jahrgangsinternen TuFTs (z.B. „Von den 13ern lerne ich nicht viel und es ist auch kein klassenübergreifendes Klima entstanden.“; „Inhalte jahrgangsspezifischer gestalten“).
- Die Fellows fordern eine stärkere Praxisorientierung der TuFTs (z.B. „Die TuFTs sollten viel strukturierter und praxisorientierter sein. Ein Beispiel was richtig gut behandelt wird aus dem Schulalltag, keine halben Sachen.“; „Der Einbau von Case-Studies oder Best-Practice Beispielen würde die TuFTs sehr bereichern.“).

#### Fortbildungen

- Die Kommentare bezüglich der Fortbildungen beschränken sich im Wesentlichen auf die terminliche Lage (z.B. „Fortbildungen nicht am Wochenende!“; „FoBis am Freitag. Der zusätzliche Samstag ist sehr kräftezehrend bei einer 50-60H-Woche. Das klingt sehr viel, trifft aber auf mehrere Fellows zu. Sicherlich wird hier an einigen Stellen freiwillig ein Mehraufwand betrieben, jedoch scheint dieser durchaus sinnvoll. Außerdem nutzen einige Fellows gerne einen Tag des Wochenendes für Unterrichtsvorbereitung, ich beobachte bei engagierten Lehrern dasselbe. Für die Trainer sind Wochenend-FoBis sicherlich auch nicht weniger anstrengend.“)
- sowie den Wunsch nach inhaltlicher Schwerpunktsetzung (z.B. „Fortbildungen auf ein Thema fokussieren, das intensiv bearbeitet wird“; „näher auf die Arbeit in verschiedenen Klassenstufen eingehen - es bleibt zu allgemein; Trennung der Fortbildungen von Sek I- und Sek II-Fellows“)
- Ein Fellow vermisst vorbereitende Literatur („Literaturhinweise zur Vorbereitung auf Fortbildungen“)

#### Hospitationen/Bilanzgespräche

- Mehrere Fellows wünschen sich häufigere Hospitationen, generell mehr Feedback und mehr Zeit für Bilanz-Gespräche – auch außerhalb der Schule (z.B. „Mehr Zeit für strukturierte Ziel-/Bilanz-Gespräche außerhalb der Schule“; „Noch intensivere Zusammenarbeit mit den Trainer\_innen“)

### 7.1.15.3 Schuleinsatz

- Einige Fellows fordern eine strukturiertere Gestaltung des Schuleinsatzstartes im Sinne einer besseren Einführung des Fellows an der Schule durch Teach First (z.B. „bessere Kommunikation/ Cooperation mit Schulleitung“; „MEHR Unterstützung in der Entwicklung des Einsatzprofils mit der Schulleitung“; „Mehr Koordination bei der

Erstellung des Stundenplans. Klarer Vorgaben, welche Klassen man wie übernehmen soll. Dies sollte auch definitiver mit der Schulleitung besprochen werden.“

- Ein Fellow macht den Vorschlag Projektideen schon während der Sommerakademie auszuarbeiten um diese schneller in der Schule umsetzen zu können („Fokus auf effektive Zeitnutzung. Wenn die Schule schon vorher das Projekt des Fellows kennen würde und ihr "go" gibt, könnte der Fellow das Projekt gleich durchführen.“)
- Ein Fellow schlägt vor den Fellows zum Start des Schuleinsatzes einen Reader mit den wichtigsten Inhalten an die Hand zu geben („Quickstart guide" zu Beginn des Schuleinsatzes, evtl. im Rahmen der ersten Fortbildung, mit den wichtigsten Inhalten der Sommerakademie, die gerade in den ersten Wochen wichtig sind.“)

#### **7.1.15.4 Professionelle Kooperation**

- Etliche Fellows äußern den Wunsch nach stärkerem bzw. intensiverem Austausch untereinander (z.B. „Vernetzung unter den Fellows: es sollten Freiräume geschaffen werden, in denen sich Fellows regional vernetzen und beispielsweise ganz konkret gemeinsam Unterricht planen können, um so effektiver und zeitsparender zu arbeiten und sich gegenseitig zu befruchten. Dafür könnten TuFTs nur noch alle zwei Monate oder mit freiwilliger Teilnahme stattfinden bzw. Zeit auf den FoBi's geschaffen werden.“; „Ich würde einen Planungsstammtisch begrüßen. Ein Tag, alle 2 Wochen, an dem man im Büro vorbeikommen und gemeinsam an der eigenen Planung feilen kann. So können die Fellows sich gegenseitig bestärken und ihre jeweiligen Erfahrungen für andere nutzbar machen. Es darf aber kein verpflichtendes Angebot sein, sondern eine freiwillige Unterrichtsplanungswerkstatt. Auf diese Art und Weise können insbesondere Planungen, die sich zunächst als schwierig herausstellen, durch Unterstützung von außen, doch bearbeitet werden.“; „intensivere Balkontreffen“; „Austausch zwischen Fellows verstärken (am besten 2 Fellows an einer Schule)“)
- Ein Fellow wünscht sich zusätzlich mehr Austausch mit den Lehrenden der Einsatzschule („Ein besserer Zugang zu Unterrichtsmaterialien und ein größerer Austausch darüber mit den LehrerInnen.“)

#### **7.1.15.5 Zertifizierung der Fellows**

- Zwei Fellows äußern den Wunsch nach Zertifizierung – einmal im Sinne von zertifizierten Angeboten, einmal im Sinne einer Möglichkeit zum Quereinstieg („Z.B: DaF /DaZ Zertifikate, Konflikttrainer, Streitschlichter usw. also nachweisbare Kompetenzen, die auch nach TFD noch anerkannt werden.“; „Begleitender Studiengang --> Möglichkeit zum Quereinstieg“).

#### **7.1.15.6 Belastungsfaktoren**

- Als belastender Faktor wird von einigen Fellows hoher Zeitdruck genannt (z.B. „Die enge Bindung an TFD schränkt mich zeitlich manchmal stark ein, sodass ich "dringende" Dinge vor "wichtigen" tue. Das tut dem Schuleinsatz nicht immer gut.“;



„Wenn ich alles so umsetzen würde, wie es sich TFD wünscht, hätte ich keine Freizeit mehr und würde nur noch für die Schule leben.“)

- Einige Fellows nehmen widersprüchliche Anforderungen wahr (z.B. „Teilweise konfliktieren die Ansprüche durch TDF mit denen durch die Schule.“) oder hohen Erwartungsdruck (z.B. „Irgendwie fühle ich mich öfters schlecht, dass ich nicht alles so umsetze, wie von TFD gewünscht, obwohl ich weiß, dass ich gute Arbeit leiste. Mir ist bewusst, dass die Erwartungen an die Fellows sehr groß sind und auch sehr hoch gesteckt sein müssen. Andererseits finde ich sie im Schulalltag unrealistisch und das erhöht den Druck bei vielen Fellows. Ob dadurch die Qualität steigt, wage ich zu bezweifeln.“).

### **Trainingsübergreifende Merkmale der Lernumgebung**

Neben den Merkmalen der einzelnen Trainingsteile, wurden auch bestimmte trainingsübergreifende Merkmale erhoben. Dazu zählen die Einschätzung der Qualität der Trainer und Trainerinnen und die Kohärenz des Programms.

#### **7.1.16 Einschätzung der Qualität der Trainer und Trainerinnen**

Zur Einschätzung der Qualität der Trainer und Trainerinnen wurden zum ersten Messzeitpunkt (3. Welle) 12 Items aus der PEB-Sem Studie übernommen. Für den zweiten Messzeitpunkt (4. Welle) wurden die bestehenden Items etwas verdichtet und zusätzliche Items zur Unterstützungsqualität während des Schuleinsatzes entwickelt. Insgesamt zeigt sich für die Aussagen, die zu zwei verschiedenen Zeitpunkten vorgelegt wurden, eine hohe Stabilität.

Die Qualität der Trainer und Trainerinnen wird von den befragten Fellows zum ersten Messzeitpunkt sehr positiv eingeschätzt. Besonders die Items zur Schülerorientierung wurden im Mittel sehr hoch eingeschätzt (train1\_3, train4\_3, train10\_3).

Aber auch die Unterstützung der Fellows durch die Trainer und Trainerinnen wird von den Befragten überdurchschnittlich hoch bewertet (train2\_3, train6\_3, train7\_3, train8\_3, train9\_3). So *trifft* im Durchschnitt *zu* bzw. *eher zu*, dass die Trainer und Trainerinnen bei Problemen mit den Fellows nach einer Lösung suchen, Ziele für die weitere pädagogische Entwicklung der Fellows formulieren und die Fellows durch ihre Beratung unterstützen.

Zum zweiten Messzeitpunkt (vierte Erhebungswelle) wurden die Items mit den höchsten Faktorladungen des ersten Messzeitpunktes beibehalten (8 Items) und zusätzlich sechs neue Items entwickelt, welche die Unterstützungsqualität während des Schuleinsatzes erfassen sollten. Die Traineritems, die schon in der zweiten Erhebungswelle verwendet wurden, werden von den Fellows in der vierten Erhebungswelle im Durchschnitt sehr ähnlich eingeschätzt. Teilweise ergeben sich identische Mittelwerte.



Variable	Einschätzung der Trainer/Trainerinnen	MZP 1 (Welle 3)		MZP 2 (Welle 4)	
		M	STD	M	STD
train1_3/ train1_4	Meine Trainer /- innen können sich gut in die Perspektive von Schülern / Schülerinnen hineinversetzen.	3,57	0,62	3,48	0,71
train2_3/ train2_4	Meine Trainer /-innen beziehen sich auf klare Kriterien für gute Lehrarbeit.	3,44	0,62	3,45	0,60
train3_3	Wenn ich den Trainern /-innen ein Problem schildere, suchen sie mit mir nach einer Lösung.	3,52	0,67	-	-
train4_3/ train4_4	Meine Trainer /-innen vermitteln Begeisterung für die Arbeit mit Kindern und Jugendlichen.	3,66	0,66	3,66	0,51
train5_3	Meine Trainer /-innen fördern die professionelle Kooperation.	3,45	0,67	-	-
train6_3/ train6_4	Die Beratungen mit den Trainern /-innen sind eine wesentliche Unterstützung für mich.	3,52	0,72	3,50	0,60
train7_3/ train7_4	Meine Trainer /-innen formulieren Ziele für meine weitere pädagogische Entwicklung.	3,28	0,78	3,34	0,64
train8_3/ train8_4	Meine Trainer/-innen begründen, warum sie bestimmte Entscheidungen treffen.	3,16	0,84	3,27	0,75
train9_3/ train9_4	Ich kann von meinen Trainer/-innen lernen, wie ich die Unterrichtszeit effektiv nutze.	3,34	0,75	3,30	0,76
train10_3/ train10_4	Ich kann von meinen Trainer/-innen lernen, wie Schüler/-innen zu motivieren sind.	3,36	0,78	3,23	0,54
train11_3	Meine Trainer/-innen sind an Fragen der Beurteilung von Schüler/-innen besonders interessiert.	2,40	0,90	-	-
train12_3	Meine Trainer/-innen sind an Fragen der Beurteilung von Unterricht besonders interessiert.	3,04	0,79	-	-
train13_4	Ich fühle mich durch meine/n TrainerIn zu wenig bei der Entwicklung eines fokussierten Einsatzprofils unterstützt.			1,96	0,81
train14_4	MeinE TrainerIn ist unverlässlich bei Absprachen.			1,45	0,91
train15_4	MeinE TrainerIn gibt mir wenig konstruktives Feedback.			1,48	0,79
train16_4	MeinE TrainerIn verweist auf Handlungsmuster, die ich in Konfliktsituationen anwenden kann.			2,98	0,80
train17_4	MeinE TrainerIn stärkt mich darin, Herausforderungen anzunehmen.			3,25	0,69
train18_4	MeinE TrainerIn unterstützt mich darin, Verantwortung für Veränderung zu übernehmen.			3,29	0,68
N		62		48	

Tabelle 21: Einschätzung der Trainer und Trainerinnen - Deskriptive Statistiken

Die Items zur Unterstützungsqualität werden insgesamt sehr positiv eingeschätzt mit Mittelwerten, die bei fast allen Items deutlich über 3,0 liegen. Hier ist zu beachten, dass die Items train13\_4, train14\_4 und train15\_4 negativ formuliert sind und daher niedrige Werte positiv zu interpretieren sind. So lehnen die Fellows im Durchschnitt die Aussagen ab, dass die Trainer und Trainerinnen sie zu wenig bei der Entwicklung eines fokussierten

Einsatzprofils unterstützen, zu wenig konstruktives Feedback geben und wenig verlässlich bei Absprachen sind (Mittelwerte: 1,96; 1,48; 1,45). Die Fellows stimmen dagegen den Aussagen, dass die Trainer und Trainerinnen sie darin stärken Herausforderungen anzunehmen und sie dabei unterstützen Verantwortung für Veränderung zu übernehmen, stark zu (Mittelwerte: 3,25; 3,29). Der Aussage, dass die Trainer und Trainerinnen auf Handlungsmuster verweisen, die in Konfliktsituationen anwendbar sind, schätzen die Fellows von allen Aussagen am schlechtesten, aber immer noch als eher zutreffend (Mittelwert: 2,98), ein.

Die Faktorenanalyse zum ersten Messzeitpunkt ergibt eine dreifaktorielle Lösung, wobei auf den dritten Faktor nur die beiden letzten Items (train11\_3 und train12\_3) laden. Auf den ersten Faktor laden die Items train1\_3, train3\_3, train4\_3, train5\_3 und train10\_3. Item train1\_3, train4\_3 und train10\_3 haben allerdings die höchsten Ladungen. Diese Items fokussieren auf die Schülerorientierung der Trainer und Trainerinnen. Auf den zweiten Faktor laden die Items train2\_3, train6\_3, train7\_3, train8\_3, und train9\_3. Diese Items fokussieren auf die Unterstützung der Fellows durch die Trainer und Trainerinnen. Die Reliabilitätsanalyse der Items des ersten Faktors ergibt ein Cronbach's alpha von 0,83, der Mittelwert der neugebildeten Skala *Schülerorientierung der Trainer* beträgt 3,52. Die Reliabilitätsanalyse der Items des zweiten Faktors ergibt ein Cronbach's alpha von 0,77. Der Mittelwert der neugebildeten Skala *Unterstützung der Trainer* beträgt hier 3,35.

Die Faktorenanalyse der Items zum zweiten Messzeitpunkt ergibt eine vierfaktorielle Lösung, wobei alle Items bis auf die neu entwickelten negativ formulierten Items mit deutlich höherer Ladung auf den ersten Faktor laden. Auf den zweiten Faktor laden nur die negativ formulierten Items train14\_4 und train15\_4 hoch. Der Screeplot spricht für eine zweifaktorielle Lösung. Das heißt die entwickelten Items zur Unterstützungsqualität während des Schuleinsatzes bilden keine eigene Dimension ab. Die anschließende Reliabilitätsanalyse ergibt für die Items des ersten Faktors eine gute Reliabilität (Cronbach's alpha: 0,87), doch für den zweiten Faktor lässt sich aufgrund der geringen Anzahl an Items keine gute Reliabilität berechnen. Der Mittelwert der neugebildeten Skala *Qualität der Trainer* beträgt 3,37.

#### 7.1.17 Kohärenz in der Ausbildung

Die Kohärenz der Ausbildung wird von den teilnehmenden Fellows im Durchschnitt sehr positiv bewertet. Die Mittelwerte der Einzelitems betragen alle 3,00 und höher. Die befragten Fellows geben an, dass die einzelnen Teile der Ausbildung gut aufeinander bezogen sind und aufeinander aufbauen. Des Weiteren nehmen sie eine ausreichende Kontinuität in der Begleitung während der Ausbildung wahr, sowie eine gute inhaltliche Absprache zwischen den für die Ausbildung Verantwortlichen. Im Mittel geben die Befragten an, dass im Laufe der Ausbildung eher deutlich ist, an welcher Kompetenz gearbeitet wird.

Variable	Kohärenz	MZP (Welle 3)	
		M	STD
koh1_3	Die einzelnen Teile meiner „Lehrerausbildung“ sind gut aufeinander bezogen.	3,24	0,56
koh2_3	Es gibt eine ausreichende Kontinuität in der Begleitung während der „Lehrerausbildung“.	3,18	0,76
koh3_3	Die Inhalte meiner „Lehrerausbildung“ bauen aufeinander auf.	3,31	0,56
koh4_3	Es gibt eine gute inhaltliche Absprache zwischen denen, die für einzelne Teile in meiner „Lehrerausbildung“ verantwortlich sind.	2,88	0,69
koh5_3	Im Laufe der „Lehrerausbildung“ ist immer deutlich, an welcher Kompetenz gerade gearbeitet wird.	3,03	0,73
koh6_3	Die Strukturierung der „Lehrerausbildung“ wurde ausreichend vermittelt.	3,40	0,64
koh7_3	Die Strukturierung der „Lehrerausbildung“ folgt einer nachvollziehbaren Logik.	3,34	0,63
koh8_3	Die Trainer/-innen haben einen hinreichenden Überblick zur gesamten „Lehrerausbildung“ gegeben.	3,21	0,79
N		62	

Tabelle 22: Einschätzung der Kohärenz der Ausbildung – Deskriptive Statistiken

Die zugehörige Faktorenanalyse ergibt eine zweifaktorielle Lösung. Auf den ersten Faktor laden die Items koh1-koh5 und auf den zweiten Faktor die Items koh6-koh8. Bei den zugehörigen Reliabilitätsanalysen wird deutlich, dass nur die Skala zum ersten Faktor sinnvoll gebildet werden kann, da die interne Konsistenz der Items des zweiten Faktors nicht ausreichend hoch ist. Die interne Konsistenz der Items des ersten Faktors ist mit einem Cronbach's alpha von 0,79 noch zufriedenstellend. Der Mittelwert der neugebildeten Skala *Kohärenz* beträgt 3,13

## 7.2 Merkmale der Nutzung der Lernumgebung

Im Folgenden werden die Prozesse der Nutzung der Lernumgebung durch die Fellows dargestellt. Dazu wurde in der ersten Erhebung nach der Nutzung von Lernstrategien, in der zweiten Erhebung nach der Nutzung von sozialen Ressourcen und in der zweiten, dritten und vierten Erhebung nach der zur Verfügung stehenden Lernzeit gefragt. Weiterhin kamen in der vierten Erhebung Fragen zur Nutzung der TFD Werkstatt hinzu.

### 7.2.1 Nutzung von Lernstrategien

Um herauszufinden ob die Fellows beim Aneignen von neuem Wissen Lernstrategien nutzen, wurde eine Skala bestehend aus 7 Items aus der PEB-Sem Studie eingesetzt. Den Aussagen zur Nutzung von Lernstrategien wurde von den teilnehmenden Fellows im Durchschnitt sehr stark zugestimmt (Mittelwerte: 3,22-3,62). Die teilnehmenden Fellows geben an, dass sie sich konkrete Beispiele zu neu Gelerntem überlegen und versuchen, Verbindungen zu früheren Erfahrungen herzustellen, sowie neue Begriffe und Theorien auf bereits bekannte zu beziehen. Am stärksten stimmen sie der Aussage zu, dass sie neues Wissen in einem Bereich auf andere Bereiche anwenden (Mittelwert: 3,62) und versuchen in Gedanken



Neues mit ihrem Vorwissen zu verbinden (Mittelwert: 3,57). Des Weiteren geben sie an, dass sie sich zu neuen Konzepten praktische Anwendungen sowie manche neuen Sachverhalte bildlich vorstellen.

Variable	Nutzung von Lernstrategien	MZP (Welle 1)	
		M	STD
elab1_1	Ich versuche in einem Bereich neu Gelerntes auch auf andere Bereiche anzuwenden.	3,62	0,68
elab2_1	Ich überlege mir konkrete Beispiele zu neu Gelerntem.	3,22	0,74
elab3_1	Ich versuche bewusst Verbindungen zwischen früheren Erfahrungen und neu Gelerntem herzustellen.	3,54	0,69
elab4_1	Ich stelle mir manche neuen Sachverhalte bildlich vor	3,29	0,82
elab5_1	Zu neuen Konzepten stelle ich mir praktische Anwendungen vor.	3,34	0,64
elab6_1	Ich versuche in Gedanken, Neues mit dem zu verbinden, was ich schon zu einem Thema weiß.	3,57	0,66
elab7_1	Ich versuche, neue Begriffe und Theorien auf mir bereits bekannte Begriffe und Theorien zu beziehen.	3,37	0,72
N		65	

Tabelle 23: Nutzung von Lernstrategien – Deskriptive Statistiken

Eine Faktorenanalyse mit den 7 Items zur Nutzung von Lernstrategien zeigt, dass alle Items auf einen Faktor laden. Allerdings zeigt sich in der anschließenden Reliabilitätsanalyse, dass die Trennschärfe von elab4\_1 grenzwertig ist und sich Cronbach's alpha bei Ausschluss dieses Items verbessert. Daher wird die Skala *Nutzung von Lernstrategien* aus den sechs verbleibenden Items gebildet. Cronbach's alpha beträgt 0,84 und spricht für eine gute Reliabilität. Der Mittelwert der Skala beträgt 3,44. Wie bereits bei der Auswertung der Einzelitems deutlich wurde, nutzen die teilnehmenden Fellows Lernstrategien eher stark.

### 7.2.2 Professionelle Kooperation

Um die professionelle Kooperation unter den Fellows zu erfassen, wurden drei Items aus der PEB-Sem Studie übernommen.

Variable	Professionelle Kooperation	MZP (Welle 2)	
		M	STD
rsco1_2	Ich spreche mit anderen, um Frustrationen loszuwerden.	2,68	1,00
rsco5_2	Ich diskutiere über die Inhalte der Ausbildung oft mit anderen.	2,66	0,73
rsco6_2	Wenn mir etwas nicht ganz klar ist, frage ich andere um Rat.	2,82	0,93
N		65	

Tabelle 24: Professionelle Kooperation – Deskriptive Statistiken

Die Mittelwerte für die Fragen zur professionellen Kooperation liegen alle etwas über dem theoretischen Mittelwert von 2,50 herum. Die teilnehmenden Fellows beantworteten die Fragen im Mittel mit *trifft eher nicht zu* und *trifft eher zu*. Am ehesten trifft auf die teilnehmenden Fellows zu, dass sie bei Unklarheiten andere um Rat fragen (Mittelwert: 2,82). Die Faktorenanalyse der drei Items zur professionellen Kooperation zeigt, dass alle



Items auf einen Faktor laden. Es handelt sich hierbei jedoch um eine reduzierte Auswahl an Items, da die weiteren Items der Skala sich auf den schulischen Kontext beziehen. Die komplette Skala sollte ursprünglich zum vierten Messzeitpunkt erhoben werden. Die Reliabilitätsanalyse ergibt allerdings keinen ausreichenden Wert für Cronbach's alpha (0,71), sodass keine Skala zur professionellen Kooperation gebildet werden kann. Aus diesem Grund und auch aus Kapazitätsgründen wurde auf eine zweite Erhebung dieser Items verzichtet.

### 7.2.3 Zeitmanagement

Sowohl zum zweiten als auch zum dritten und vierten Messzeitpunkt wurden die teilnehmenden Fellows zu ihrem Zeitmanagement befragt. Es wurde die Skala zum Zeitmanagement, bestehend aus drei Items, aus der PEB-Sem Studie eingesetzt<sup>7</sup>. Die Items sind alle negativ formuliert, sodass hier kleinere Werte positiv zu interpretieren sind. Zum Zeitpunkt der zweiten Befragung lag der Fokus auf der Arbeit im Rahmen des Online Campus, während die Fragen zum Zeitpunkt der dritten und vierten Befragung auf Seminarveranstaltungen und Unterricht ausgerichtet waren.

Variable	Zeitmanagement Online Campus	MZP (Welle 2)	
		M	STD
time1_2	Es fällt mir schwer, genügend Zeit zur Bearbeitung der Reader des Online Campus aufzubringen.	2,54	1,02
time2_2	Es fällt mir schwer, genügend Zeit zur Bearbeitung der Portfolio-Aufgaben des Online Campus aufzubringen.	2,58	1,01
time3_2	Für Entspannung und Freizeit bleibt mir letztlich keine Zeit.	2,23	1,07
N		65	

Tabelle 25: Zeitmanagement Online Campus - Deskriptive Statistiken

Variable	Zeitmanagement Seminarteil/MSact!/Schuleinsatz	MZP 1 (Welle 3)		MZP 2 (Welle 4)	
		M	STD	M	STD
time1_3/ time1_4	Es fällt mir schwer, genügend Zeit zur Vor- und Nachbereitung von Veranstaltungen aufzubringen.	2,35	0,89	2,56	0,85
time2_3/ time2_4	Es fällt mir schwer, genügend Zeit zur Vor- und Nachbereitung von Unterricht aufzubringen.	2,50	0,95	2,71	0,82
time3_3/ time3_4	Für Entspannung und Freizeit bleibt mir letztlich keine Zeit.	2,60	0,82	2,46	0,80
N		62		48	

Tabelle 26: Zeitmanagement Seminarteil/ MSact!/ Schuleinsatz - Deskriptive Statistiken

Im Wesentlichen unterscheiden sich die Einschätzungen der teilnehmenden Fellows zu ihrem Zeitmanagement nicht zwischen den drei Messzeitpunkten. Alle Mittelwerte liegen um den theoretischen Mittelwert von 2,50 herum, d.h. die Fellows nehmen weder einen großen Zeitmangel noch einen großen Überfluss an Zeit wahr. Lediglich Item time2\_4 ist

<sup>7</sup> Diese wurden für die erste Erhebung leicht adaptiert und für die dritte und vierte Erhebung in der Originalfassung übernommen.

leicht gestiegen (Mittelwert: 2,72). Für die Items zum Zeitmanagement im Rahmen des Online Campus ergibt sich in der Faktorenanalyse eine einfaktorielle Lösung. Aus der Reliabilitätsanalyse erhält man ein Cronbach's alpha von 0,87, welches für eine gute Reliabilität spricht. Die Skala *Zeitmanagement Online Campus* hat einen Mittelwert von 2,45 und liegt damit leicht unter dem theoretischen Mittelwert. Für die Items zum zweiten Messzeitpunkt ergibt sich ebenfalls eine einfaktorielle Lösung, allerdings zeigt die Reliabilitätsanalyse eine unzureichende Reliabilität der Items an. Dies liegt an einer schlechten Trennschärfe von time3\_3. Daher ist hier keine Skalenbildung möglich. Gleiches gilt für den dritten Messzeitpunkt. Auch hier ergibt sich eine einfaktorielle Lösung, aber keine hinreichende Reliabilität der Skala.

#### 7.2.4 Nutzung der TFD Werkstatt

Zur Evaluation der Nutzung der TFD Werkstatt durch die Fellows wurden drei Items entwickelt. Die Fellows schätzen die Nutzung der Werkstatt im Durchschnitt eher negativ ein. Sie geben an, dass sie eher keine eigenen Materialien hochgeladen (Mittelwert: 1,65) und auch eher noch keine Materialien von anderen zur Nutzung heruntergeladen haben (Mittelwert: 1,77). Dafür lehnen sie die Aussage eher ab, dass sie wenig relevante Informationen auf den Seiten der Werkstatt gefunden haben (Mittelwert: 2,58). Das heißt sie finden zwar Informationen, nutzen diese aber eher weniger.

Variable	Nutzung der TFD Werkstatt	MZP (Welle 4)	
		M	STD
werkstatt8_4	Auf den Seiten der Werkstatt habe ich bisher wenig für mich relevante Informationen gefunden.	2,58	0,90
werkstatt9_4	Ich habe bereits von mir entworfene Materialien (Arbeitsblätter, Stundenentwürfe, Projektpläne, o.ä.) auf den Seiten der Werkstatt hochgeladen, so dass andere sie nutzen können.	1,65	1,02
werkstatt10_4	Ich habe bereits Materialien für meine Arbeit genutzt, die andere Fellows auf den Seiten der Werkstatt hochgeladen hatten.	1,77	0,86
N		48	

Tabelle 27: Nutzung der TFD Werkstatt - Deskriptive Statistiken

Bei der Faktorenanalyse ergibt sich eine einfaktorielle Lösung. Die anschließende Reliabilitätsanalyse zeigt allerdings, dass die Items werkstatt8\_4 und werkstatt9\_4 eine sehr schlechte Trennschärfe aufzeigen. Somit ergibt sich keine zufriedenstellende Reliabilität (Cronbach's alpha: 0,41) und es ist keine Skalenbildung möglich.

## 8 Ergebnisse

Daten zu den Ergebnissen des Trainingsprogramms wurden in Form von Selbstwirksamkeitserwartungen, selbsteingeschätzten Unterrichtskompetenzen und Testergebnissen aus einem Wissenstest erhoben. Sämtliche Ergebnisse wurden dreimal erhoben, um den Zuwachs an Kompetenz und Wissen untersuchen zu können. In Welle 1, Welle 2 und Welle 4 wurde ein umfangreicher Wissenstest durchgeführt sowie in Welle 3 einige Wissensfragen zum Themenbereich Leadership gestellt. Die Selbstwirksamkeitserwartungen und selbsteingeschätzte Unterrichtskompetenz wurden in der ersten, dritten und vierten Erhebung erfasst. Da wir in den einzelnen Erhebungswellen unterschiedlich große Stichproben haben, basieren die Analysen auf unterschiedlichen Fallzahlen.

### 8.1 Selbstwirksamkeitserwartung

Die Items der Selbstwirksamkeitserwartung sind von Tschannen-Moran und Woolfolk-Hoy (2001) übernommen und wurden vom Forscherteam aus dem Englischen übersetzt. Im englischen Original gibt es jeweils 8 Items zu den Aspekten Klassenführung, Unterrichtsstrategien, sowie Schülermotivation. Die Items mussten auf einer Skala von 1 – *keine Fähigkeit* bis 9 – *sehr große Fähigkeit* eingeschätzt werden. Insgesamt schätzen die teilnehmenden Fellows zum ersten Erhebungszeitpunkt – also bevor das Training begonnen hat – ihre Selbstwirksamkeit zu den Aspekten Klassenführung, Unterrichtsstrategien und Schülermotivation überdurchschnittlich hoch ein. Die Item-Mittelwerte liegen alle über dem theoretischen Mittelwert von 5,00. Besonders hoch schätzen die Befragten im Durchschnitt ihre Fähigkeiten in den Bereichen Unterrichtsstrategien und Schülermotivation ein. Im Bereich Unterrichtsstrategien liegen die Mittelwerte aller Items deutlich über 6,00 mit Ausnahme der Fähigkeit verschiedene Beurteilungsmethoden anzuwenden. Hier liegt der Mittelwert mit 5,92 knapp darunter. Im Bereich Schülermotivation sehen die Einschätzungen im Durchschnitt sehr ähnlich aus. Hier liegen alle Item-Mittelwerte um 6,50 herum mit Ausnahme der Fähigkeit auch die schwierigsten Schüler und Schülerinnen zu erreichen (Mittelwert: 5,92). Die Einschätzungen zum zweiten Messzeitpunkt (3. Erhebungswelle) sehen im Wesentlichen ähnlich aus wie zum ersten Messzeitpunkt. Mit Ausnahme der Fähigkeiten, verschiedene Beurteilungsmethoden anzuwenden und Familien zu unterstützen, damit sie ihren Kindern helfen können, schätzen die Fellows ihre Fähigkeiten wieder alle über dem theoretischen Mittelwert von 5,00 ein.

Variable	Selbstwirksamkeitserwartung — Klassenführung	MZIP 1 (Welle 1)		MZIP 2 (Welle 3)		MZIP 3 (Welle 4)	
		M	STD	M	STD	M	STD
	Einleitungstext: Wie hoch schätzen Sie Ihre Fähigkeit ein, auch in schwierigen Schulsituationen...						
seff9_1/ seff9_3/ seff9_4	... störendes Verhalten im Klassenzimmer zu kontrollieren?	5,48	1,38	6,00	1,67	5,14	1,35
seff10_1/ seff10_3/ seff10_4	... Schüler/innen dazu zu bewegen, sich an die Klassenregeln halten?	6,02	1,16	6,05	1,41	5,33	1,39
seff11_1/ seff11_3/ seff11_4	... einen Schüler / eine Schülerin zu beruhigen, der/die laut ist oder stört?	5,52	1,21	5,76	1,53	5,48	1,23
seff12_1/ seff12_3/ seff12_4	... mit den Schülern und Schülerinnen ein Regelsystem für die Klasse zu etablieren?	6,60	1,15	5,81	1,71	5,29	1,57
seff13_1/ seff13_3/ seff13_4	... zu verhindern, dass wenige Problemschüler/innen eine komplette Unterrichtsstunde ruinieren?	5,26	1,50	5,10	1,53	5,05	1,36
seff14_1/ seff14_3/ seff14_4	... mit aufsässigen Schülern und Schülerinnen umzugehen?	5,40	1,42	5,38	1,72	5,43	1,50
seff15_1/ seff15_3/ seff15_4	... ihre Erwartungen an Schülerverhalten klar zu formulieren?	6,24	1,21	6,10	1,48	5,95	1,34
seff16_1/ seff16_3/ seff16_4	... Routinen einzusetzen, um reibungslose Abläufe zu gewährleisten?	6,31	1,30	5,48	1,58	5,57	1,35
N		42					

Tabelle 28: Selbstwirksamkeitserwartung - Klassenführung - Deskriptive Statistiken

Die Fähigkeit, Familien zu unterstützen, wird mit einem Mittelwert von 3,24 deutlich schlechter eingeschätzt als alle anderen Fähigkeiten. Dies könnte damit zusammenhängen, dass sich im Laufe der Qualifizierung herausstellt, dass diese Fähigkeit nicht zum Einsatzprofil der Fellows gehört. Die Fähigkeit, störendes Verhalten im Klassenzimmer zu kontrollieren wird, bzw. störende Schüler zu beruhigen, wird zum zweiten Messzeitpunkt etwas besser eingeschätzt (Mittelwerte: 6,00; 5,76). Insgesamt ist allerdings zu beobachten, dass sich die Fellows im Durchschnitt etwas schlechter einschätzen als zum ersten Messzeitpunkt. Die Tatsache, dass sich die teilnehmenden Fellows zum zweiten Messzeitpunkt schlechter einschätzen, könnte damit zusammenhängen, dass sie während der Sommerschule erste Erfahrungen im Unterrichten sammeln konnten und ihre Einschätzungen dadurch realistischer geworden sind. Zum dritten Messzeitpunkt, der nach Beginn des Schuleinsatzes lag, schätzen die Fellows ihre Selbstwirksamkeit im Mittel noch immer überdurchschnittlich ein. Die beiden Ausnahmen bilden die gleichen Items wie bereits zum zweiten Messzeitpunkt (s.o.: Bewertungsmethoden; Familien unterstützen).



Variable	Selbstwirksamkeitserwartung — Unterrichtsstrategien	MZP 1 (Welle 1)		MZP 2 (Welle 3)		MZP 3 (Welle 4)	
		M	STD	M	STD	M	STD
	<b>Einleitungstext:</b> Wie hoch schätzen Sie Ihre Fähigkeit ein, auch in schwierigen Schulsituationen...						
seff1_1/ seff1_3/ seff1_4	... verschiedene Bewertungsmethoden im Unterricht anzuwenden?	5,67	1,83	4,43	1,42	4,67	1,24
seff2_1/ seff2_3/ seff2_4	... eine alternative Erklärung oder ein weiteres Beispiel anzuführen, wenn Ihre Schüler/innen Verständnisprobleme haben?	7,50	1,27	7,00	1,59	6,95	1,50
seff3_1/ seff3_3/ seff3_4	... gute Fragen für Ihre Schüler/innen zu konzipieren?	6,55	1,37	6,00	1,61	5,76	1,16
seff4_1/ seff4_3/ seff4_4	... alternative Unterrichtsverfahren im Klassenzimmer einzusetzen?	6,62	1,62	5,73	1,78	5,33	1,52
seff5_1/ seff5_3/ seff5_4	... auf komplizierte Fragen Ihrer Schüler/innen zu antworten?	6,95	1,23	6,19	1,47	6,38	1,43
seff6_1/ seff6_3/ seff6_4	... ihre Unterrichtsstunden auf ein für die einzelnen Schüler/innen angemessenes Niveau anzupassen?	6,36	1,36	5,48	1,64	5,52	1,17
seff7_1/ seff7_3/ seff7_4	... zu beurteilen, wie viel die Schüler/innen von dem verstanden haben, was Sie gelehrt haben?	6,26	1,40	5,29	1,44	5,57	1,27
seff8_1/ seff8_3/ seff8_4	... angemessene Herausforderungen für sehr begabte Schüler(innen) zu schaffen?	6,52	1,57	5,29	1,80	5,19	1,38
N		42					

Tabelle 29: Selbstwirksamkeitserwartung - Unterrichtsstrategien - Deskriptive Statistiken

Allerdings setzt sich der Trend der schlechteren oder realistischeren Einschätzung vom ersten zum zweiten Messzeitpunkt, vom zweiten zum dritten Messzeitpunkt weiter fort. Die Einschätzungen der Fellows zum dritten Messzeitpunkt sind für die Mehrzahl der Items schlechter als zum zweiten Messzeitpunkt. Dies ist besonders gut bei den Items zur Schülermotivation zu erkennen. Hier wird lediglich die Fähigkeit, Familien zu unterstützen, leicht höher eingeschätzt als zum zweiten Messzeitpunkt. Gleiches gilt für die Items zur Klassenführung. Hier werden alle Items bis auf zwei schlechter eingeschätzt als zum zweiten Messzeitpunkt: Fellows nehmen sich als routinierter wahr und schätzen sich etwas besser darin ein, mit aufsässigen Schülern und Schülerinnen umgehen zu können. Bei den Items zu Unterrichtsstrategien lässt sich der Trend am schlechtesten erkennen. Hier wird die Hälfte leicht schlechter bzw. in etwa gleich eingeschätzt.

Variable	Selbstwirksamkeitserwartung — Schülermotivation	MZP 1 (Welle 1)		MZP 2 (Welle 3)		MZP 3 (Welle 4)	
		M	STD	M	STD	M	STD
	<b>Einleitungstext:</b> Wie hoch schätzen Sie Ihre Fähigkeit ein, auch in schwierigen Schulsituationen...						
seff17_1/ seff17_3/ seff17_4	... Schüler/innen davon zu überzeugen, dass sie gut in der Schule sein können?	7,40	1,17	6,85	1,37	6,19	1,66
seff18_1/ seff18_3/ seff18_4	... den Schüler/innen dabei zu helfen, Lernen wertzuschätzen?	6,36	1,39	6,02	1,85	5,14	1,20
seff19_1/ seff19_3/ seff19_4	... Schüler/innen zu motivieren, die geringes Interesse am Arbeiten in der Schule haben?	6,55	1,13	6,19	1,53	5,38	1,34
seff20_1/ seff20_3/ seff20_4	... Familien zu unterstützen, damit sie ihren Kindern helfen können, gut in der Schule zu sein?	6,24	1,48	3,24	1,50	3,38	1,67
seff21_1/ seff21_3/ seff21_4	... das Lernverständnis von leistungsschwachen Schüler/innen zu verbessern?	6,26	1,40	5,49	1,40	5,33	1,32
seff22_1/ seff22_3/ seff22_4	... Schüler(innen) zu unterstützen, kritisch zu denken?	7,52	1,25	6,46	1,42	5,81	1,66
seff23_1/ seff23_3/ seff23_4	... die Kreativität von Schülern und Schülerinnen zu fördern?	6,74	1,43	6,05	1,61	5,29	1,44
seff24_1/ seff24_3/ seff24_4	... die schwierigsten Schüler/innen zu erreichen?	5,74	1,65	5,71	1,92	5,48	1,64
N		42					

Tabelle 30: Selbstwirksamkeitserwartung - Schülermotivation - Deskriptive Statistiken

Die Faktorenanalyse zeigt für die erste Erhebungswelle nicht wie erwartet eine dreifaktorielle, sondern eine sechsfaktorielle Struktur. Auf Grundlage des Screeplots lässt sich eine vierfaktorielle Struktur rechtfertigen, wobei auf den 4. Faktor nur drei Items laden (zwei Items zu Klassenführung und eins zu Unterrichtsstrategien). Die Überprüfung der Reliabilität der drei Items ergibt keinen akzeptablen Wert für die interne Konsistenz der Items, sodass hier von einer Skalenbildung abgesehen wird. Auf den ersten Faktor laden sechs Items zur Klassenführung, auf den zweiten Faktor sechs Items zu Unterrichtsstrategien und auf den dritten Faktor 5 Items zu Schülermotivation. Die Reliabilitätsanalysen für die drei Faktoren ergeben für den ersten und dritten Faktor (Klassenführung bzw. Schülermotivation) einen guten bzw. zufriedenstellenden Wert für die interne Konsistenz (Cronbach's alpha: 0,86 bzw. 0,80) und für den zweiten Faktor (Unterrichtsstrategien) einen akzeptablen Wert von 0,75. Für die zweite Erhebung lassen sich die Ergebnisse der Faktorenanalyse nicht replizieren. Insgesamt laden deutlich weniger Items, die als theoretisch zugehörig zu einem Aspekt der Selbstwirksamkeitserwartung entwickelt wurden, auf einen gemeinsamen Faktor. In der vierten Erhebung ergibt sich eine siebenfaktorielle

Struktur. Mit Ausnahme der Items zur Klassenführung, die das gleiche Ladungsmuster zeigen wie in der ersten Erhebung (die ersten sechs Items laden auf einen gemeinsamen Faktor – hier Faktor 2), zeigt die Faktorenanalyse ein sehr uneinheitliches Bild. Items unterschiedlicher theoretischer Konstrukte laden gemeinsam auf einen Faktor und die Items weisen häufig Doppelladungen auf. Aus diesem Grund wird hier nur die interne Konsistenz der Items des zweiten Faktors geprüft. Die Reliabilitätsanalyse zeigt eine zufriedenstellende interne Konsistenz der Items an (Cronbach's alpha: 0,80).

Aufgrund der uneinheitlichen Ergebnisse für die späteren Erhebungszeitpunkte beziehen wir uns für den Vergleich auf Ebene der Skalen auf die Ergebnisse der Faktoren- und Reliabilitätsanalysen aus der ersten Erhebungswelle und bilden für die folgenden Erhebungswellen die Skalen analog. Abbildung 4 zeigt einen Vergleich der Mittelwerte der Skalen Unterrichtsstrategien, Klassenführung und Schülermotivation für die 42 Personen, für die zu allen Messzeitpunkten Daten vorliegen.

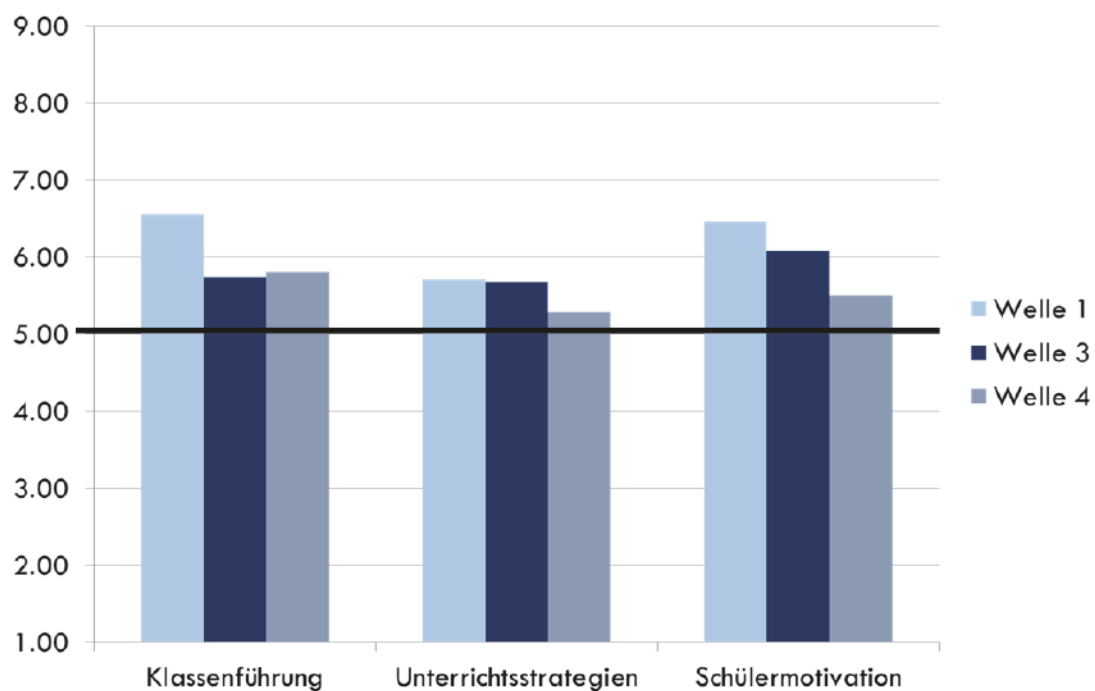


Abbildung 4: Selbstwirksamkeitserwartungen über drei Messzeitpunkte

## 8.2 Selbsteingeschätzte Unterrichtskompetenz

Die Items zur Einschätzung der Unterrichtskompetenz sind aus der PEB-Sem Studie zur Evaluation hessischer Studienseminare (Abs et al., 2009) übernommen. Die teilnehmenden Fellows waren aufgefordert ihre gegenwärtige Beherrschung für die jeweiligen Standards pädagogischen Handelns in Form von Schulnoten 1 – *ausgezeichnet*, 2 – *gut*, 3 – *befriedigend*, 4 – *ausreichend*, 5 – *nicht ausreichend* anzugeben. In diesem Fall bedeutet ein höherer Wert also eine schlechtere Einschätzung.

Variable	Standards Unterrichtshandeln	MZP 1 (Welle 1)		MZP 2 (Welle 3)		MZP3 (Welle 4)	
		M	STD	M	STD	M	STD
PEB1_1/ PEB1_3/ PEB1_4	Wie man im Unterricht an das Vorwissen der Schülerinnen und Schüler anknüpft.	3,07	1,09	3,02	0,95	2,76	0,76
PEB2_1/ PEB2_3/ PEB2_4	Wie Fehler mit Schülerinnen und Schülern so besprochen werden können, dass sie davon profitieren.	2,67	1,00	2,55	0,77	2,50	0,86
PEB5_1/ PEB5_3/ PEB5_4	Die Phasen des Unterrichts, in denen Schülerinnen und Schüler aufnehmen, verarbeiten und kontrollieren, eindeutig zu bestimmen.	3,60	1,08	2,50	0,80	2,74	0,73
PEB6_1/ PEB6_3/ PEB6_4	Wie man Unterrichtszeit effektiv nutzt.	3,29	1,02	2,52	0,77	2,86	0,61
PEB8_1/ PEB8_3/ PEB8_4	Die Wahl der Unterrichtsmethoden und Sozialformen zu begründen.	3,24	1,36	2,14	0,72	2,33	0,75
PEB10_1/ PEB10_3/ PEB10_4	Unterrichtsmaterialien selbst herzustellen.	2,62	1,15	2,45	0,86	2,24	0,76
PEB11_1/ PEB11_3/ PEB11_4	Die ausgewählten Lerninhalte logisch zu gliedern.	2,19	1,09	2,17	0,79	2,19	0,80
PEB12_1/ PEB12_3/ PEB12_4	Mich im Unterricht an meiner Planung zu orientieren und dennoch auf Unvorhergesehenes flexibel zu reagieren.	2,38	0,99	2,10	0,88	2,21	0,75
PEB14_1/ PEB14_3/ PEB14_4	Lehrmaterial auf der Basis eigener Bewertungen auszuwählen.	3,10	0,96	2,50	0,94	2,43	0,67
PEB15_1/ PEB15_3/ PEB15_4	Wie man Schülern und Schülerinnen reale Erfahrungen ermöglicht und diese mit bereits vermitteltem Wissen koppelt.	2,52	0,99	2,71	0,92	2,83	0,79
PEB19_1/ PEB19_3/ PEB19_4	Wie man Hausaufgaben sinnvoll stellen kann.	3,17	1,01	3,51	0,95	3,17	0,88
PEB20_1/ PEB20_3/ PEB20_4	Wie man Hausaufgaben sinnvoll in die weitere Unterrichtsarbeit integrieren kann.	3,12	0,94	3,56	1,00	3,29	0,86

Tabelle 31: Selbsteingeschätzte Unterrichtskompetenz über drei Messzeitpunkte - Deskriptive Statistiken

Zum ersten Messzeitpunkt schätzen die teilnehmenden Fellows ihre Kompetenzen in Bezug auf sieben von zwölf Standards durchschnittlich als befriedigend ein. Am schlechtesten wird die Kompetenz eingeschätzt, die Phasen des Unterrichts eindeutig zu bestimmen (Mittelwert: 3,60). In Bezug auf fünf von zwölf Standards schätzen sich die Fellows etwas besser als befriedigend ein. Am besten schätzen sie im Durchschnitt ihre Kompetenz ein, die ausgewählten Lerninhalte logisch zu gliedern (Mittelwert: 2,19). Im Gegensatz zu den Selbstwirksamkeitserwartungen ist bei den Kompetenzeinschätzungen vom ersten zum zweiten Messzeitpunkt (dritte Erhebung) keine Verschlechterung zu beobachten, sondern



eher eine Tendenz zur Verbesserung. Die Kompetenz, verschiedene Phasen des Unterrichts zu bestimmen, Unterrichtszeit effektiv zu nutzen, die Wahl der Unterrichtsmethoden und Sozialformen zu begründen und Lehrmaterial auf der Basis eigener Bewertung auszuwählen, wird von den Fellows im Durchschnitt deutlich besser eingeschätzt als zum ersten Messzeitpunkt. Eine leichte Verbesserung vom ersten zum zweiten Messzeitpunkt ist bei den Standards, Fehler so zu besprechen, dass die Schüler davon profitieren, Unterrichtsmaterialien selbst herzustellen und im Unterricht auf Unvorhergesehenes flexibel zu reagieren, festzustellen. Die Kompetenz bezüglich der Standards, an das Vorwissen der Schüler und Schülerinnen anzuknüpfen und die ausgewählten Lerninhalte logisch zu gliedern, wird zu beiden Messzeitpunkten in etwa gleich eingeschätzt. Die Kompetenzen bezüglich der Standards, Schülern reale Erfahrungen zu ermöglichen, Hausaufgaben sinnvoll zu stellen und sie in die Unterrichtsarbeit zu integrieren, werden hingegen zum zweiten Messzeitpunkt im Durchschnitt schlechter eingeschätzt als zum ersten Messzeitpunkt. Vom zweiten zum dritten Messzeitpunkt ist die Tendenz zur Verbesserung nicht mehr ganz so ausgeprägt wie vom ersten zum zweiten Messzeitpunkt. Insgesamt lassen sich bei fünf von zwölf Items leichte Verbesserungen feststellen. Die Kompetenzen bezüglich der Standards, Fehler so zu besprechen, dass die Schüler davon profitieren und die ausgewählten Lerninhalte logisch zu gliedern, wird zum zweiten und dritten Messzeitpunkt in etwa gleich eingeschätzt. Die Kompetenzen bezüglich der Standards, Unterrichtszeit effektiv zu nutzen, die Wahl der Unterrichtsmethoden und Sozialformen zu begründen, im Unterricht auf Unvorhergesehenes flexibel zu reagieren, Schülern reale Erfahrungen zu ermöglichen und verschiedene Phasen des Unterrichts zu bestimmen, hingegen, werden zum dritten Messzeitpunkt im Durchschnitt schlechter eingeschätzt als zum zweiten Messzeitpunkt.

Die Faktorenanalyse zeigt für die 12 Items zu Unterrichtstätigkeiten der ersten Erhebungswelle eine zweifaktorielle Struktur. Auf den ersten Faktor laden 8 Items und auf den zweiten Faktor die verbleibenden 4 Items. Der Screeplot spricht allerdings dafür, nur den ersten Faktor zu berücksichtigen. Die Überprüfung der Reliabilität der Items des ersten Faktors ergab einen sehr guten Wert für die interne Konsistenz (Cronbach's alpha: 0,91). Der Mittelwert der neugebildeten Variable *untkomp\_1* beträgt 3,03<sup>8</sup>. Für die zweite Erhebung lassen sich die Ergebnisse der Faktorenanalyse nicht replizieren. Hier ergibt sich ein sehr uneinheitliches Bild. Es wird eine vierfaktorielle Struktur angezeigt, bei der auf den ersten Faktor 2 Items, auf den zweiten Faktor 4 Items, auf den dritten Faktor 3 Items und auf den vierten Faktor 3 Items laden. Die größte gemeinsame Übereinstimmung von Items auf einem gemeinsamen Faktor zwischen erstem und zweitem Messzeitpunkt (3. Erhebung) sind drei Items (PEB\_5, PEB\_6, PEB\_10). Auch zum dritten Messzeitpunkt lässt sich in der Faktorenanalyse die Faktorenstruktur des ersten Messzeitpunktes nicht replizieren. Es wird eine fünffaktorielle Lösung angezeigt, wobei immer nur zwei bis drei Items auf einen gemeinsamen Faktor laden. Eine Skalenbildung ist aufgrund dieser Ergebnisse nicht möglich. Abbildung 5 zeigt die durchschnittliche Unterrichtskompetenz basierend auf den

---

<sup>8</sup> Dieser Mittelwert gilt für die Gesamtstichprobe der ersten Welle von 65 Fellows. In Abbildung 5 ist der Mittelwert für die reduzierte Stichprobe von 42 Fellows angegeben. Dieser liegt etwas höher.

Ergebnissen der Faktoren- und Reliabilitätsanalyse aus der ersten Erhebungswelle. D.h. hier sind nur die Items PEB1, PEB2, PEB5, PEB6, PEB8, PEB14, PEB19 und PEB20 enthalten. Außerdem ist, wie schon bei der Selbstwirksamkeit, zu beachten, dass bei der Berechnung der mittleren Unterrichtskompetenz nur die 42 Fellows einbezogen wurden, für die zu allen drei Messzeitpunkten Daten vorliegen.

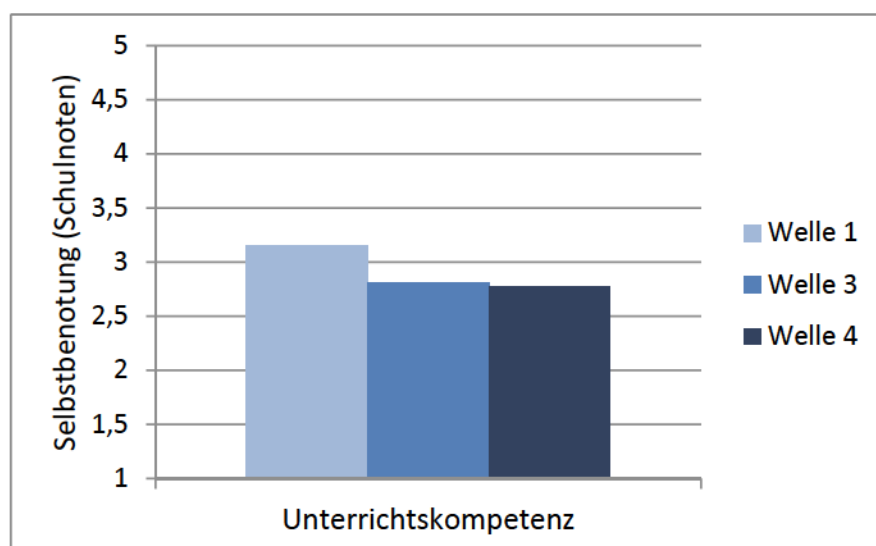


Abbildung 5: Durchschnittliche selbsteingeschätzte Unterrichtskompetenz über drei Messzeitpunkte

Zusätzlich zu den Items zur Unterrichtskompetenz, wurden aus der PEB-Sem-Studie weitere fünf Einzelitems ausgewählt, die Kompetenzen erfassen, die beim Schuleinsatz der Fellows zum Tragen kommen.

Variable	Standards Einzelitems „Schuleinsatz“	MZP 1 (Welle 1)		MZP 2 (Welle 3)		MZP3 (Welle 4)	
		M	STD	M	STD	M	STD
PEB42_1/ PEB42_3/ PEB42_4	Gruppeneinteilungen nach unterschiedlichen Kriterien und Prinzipien vorzunehmen.	2,76	1,05	2,17	1,00	2,60	0,89
PEB47_1/ PEB47_3/ PEB47_4	Wie man sich vor Überbelastung wirkungsvoll schützen kann	3,31	1,00	3,31	0,90	3,33	0,85
PEB52_1/ PEB52_3/ PEB52_4	Wie ich mit Schülerinnen und Schülern Lernstrategien erarbeiten kann.	3,40	1,01	3,07	0,91	3,29	0,81
PEB54_1/ PEB54_3/ PEB54_4	wie die Mitbestimmung von Schülerinnen und Schülern in der Schule gefördert werden kann.	2,88	1,09	2,74	1,19	2,71	0,81
PEB57_1/ PEB57_3/ PEB57_4	Wie schulische Ganztagsangebote sinnvoll für das Lernen gestaltet werden können.	2,83	1,03	2,87	0,92	3,24	1,08
N		42					

Tabelle 32: Kompetenzeinschätzung Einzelitems "Schuleinsatz" - Deskriptive Statistiken

Zum ersten Messzeitpunkt liegen die Einschätzungen der Standards im Durchschnitt um 3,0 herum. Am schlechtesten schätzen die Fellows ihre Kompetenz bezüglich des Standards, wie Lernstrategien zu erarbeiten sind, ein (Mittelwert: 3,40). Zum zweiten Messzeitpunkt sind die Einschätzungen in etwa gleich geblieben, wobei ein leichter Trend zu besseren Einschätzungen zu erkennen ist. Am stärksten ausgeprägt ist dieser für den Standard, Gruppeneinteilungen nach unterschiedlichen Kriterien und Prinzipien vorzunehmen (Mittelwert: 2,17). Auch zum dritten Messzeitpunkt ergibt sich ein ähnliches Bild. Im Durchschnitt schätzen die Fellows ihre Kompetenzen bezüglich der Standards erneut um 3,0 herum ein. Im Vergleich zum zweiten Messzeitpunkt ist nun eher ein Trend zu schlechteren Einschätzungen erkennbar. Die deutlichste Verschlechterung in der Einschätzung ist am Standard, Gruppeneinteilungen nach unterschiedlichen Kriterien und Prinzipien vorzunehmen (Mittelwert: 2,60), zu erkennen. Die Einschätzung zum dritten Messzeitpunkt ist damit fast genauso schlecht wie zu Beginn (erster Messzeitpunkt), aber insgesamt trotzdem der Standard, für den die Fellows ihre Kompetenz am besten einschätzen. Eine Faktorenanalyse wurde mit den Items nicht gerechnet, da es sich explizit um ausgewählte Einzelitems handelt und nicht um Items, die ein theoretisches Konstrukt abbilden sollen.

### 8.3 Testergebnisse

Teil der ersten und zweiten Befragung war ein umfangreicher Test, der auf die Lerninhalte des TFD Trainingsprogramms abzielte. Dabei stellt der erste Messzeitpunkt die Baseline des Wissens der Befragten dar, denn der Test fand vor der Online Campus Lernphase statt. Die teilnehmenden Fellows wurden zu Beginn der Erhebung gefragt, wie vertraut Sie bereits mit dem Online Campus sind, um eventuelles Vorwissen zu kontrollieren. Dabei geben 86,2% der Teilnehmer an, noch nie mit dem Online Campus gearbeitet zu haben, 10,8%, nur ein bisschen durchgescrollt zu haben und 2 Personen geben an, bisher eine Sitzung bearbeitet zu haben. Zum dritten Messzeitpunkt wurde auf den ausführlichen Test verzichtet, dafür aber weitere fünf Testfragen zum Bereich Leadership gestellt. Für die vierte Erhebung wurde erneut der ausführliche Test gestellt, allerdings in leicht reduzierter Variante: es wurden drei Items aus dem ursprünglichen Test gestrichen und zwar diese, die schon zum ersten Messzeitpunkt von mindestens 90% der Befragten richtig beantwortet wurden. Gleichzeitig wurden die fünf Fragen aus der dritten Erhebung in den Test integriert.

Die Fragen des Tests wurden zum einen auf Grundlage der TFD Lehrmaterialien vom Forscherteam konzipiert, zum anderen wurden inhaltlich verwandte Fragen aus dem Teilttest Pädagogisches Unterrichtswissen aus der TEDS-M Studie ausgewählt, um dadurch den Vergleich mit einer unabhängigen Stichprobe zu ermöglichen. Die Testfragen umfassen die Themenbereiche: Unterrichten (Planen sowie Umsetzen), Diagnose, Sprachförderung, Leadership und TFD-Einsatz. Insgesamt umfasst der ursprüngliche Test aus den ersten beiden Erhebungswellen 23 Testfragen, die allerdings teilweise aus mehreren Unterfragen bestehen, sodass sich hier insgesamt 52 Testitems ergeben. Von den 52 Testitems haben 26 Items ein geschlossenes Multiple Choice Antwortformat. 8 Fragen sind aus der TEDS-M



Studie übernommen. Abbildung 6 zeigt ein Beispiel für ein Multiple Choice Testitem aus dem Themenbereich Leadership aus der 2. Woche des Online Campus.

**In Deutschland korrelieren verschiedene Personenmerkmale statistisch mit geringeren Bildungschancen. Für welches der folgenden Merkmale trifft das NICHT zu?**

- ☐ Für Mädchen.
- ☐ Für Migranten.
- ☐ Für Kinder/Jugendliche mit Eltern ohne Abitur.
- ☐ Für Kinder/Jugendliche aus Elternhäusern mit geringem Einkommen.
- ☐ Ich weiß es nicht.

Abbildung 6: Beispieltestitem zu Lerninhalten aus Woche 2 des Online Campus

Für den Test der vierten Erhebung ergibt sich, zum einen aufgrund der Streichungen, zum anderen aufgrund der Integration der Testfragen aus der dritten Erhebung, eine etwas andere Anzahl an Items. Der Test setzt sich aus den 8 TEDS-M Items zum Pädagogischen Unterrichtswissen und 17 Items zu TFD Inhalten zusammen, sodass sich unter Berücksichtigung der Unterfragen 55 Items ergeben.

Die folgenden Tabellen zeigen die Testergebnisse der Antworten für die drei Messzeitpunkte nach Themenbereichen getrennt. Es werden jeweils die absoluten und relativen Häufigkeiten für die richtigen und falschen bzw. nicht gewussten Antworten berichtet. Für Fragen, die im Test aus mehreren Items bestehen, sind in der Tabelle die Zahlen insofern zusammengefasst, als dass angegeben wird, wie viele Personen die gesamte Frage richtig bzw. falsch beantwortet haben. Ausführliche Tabellen mit Aufschlüsselung der Zahlen für die Unterfragen finden sich im Anhang.

Die Stichprobe setzt sich aus 46 Fellows zusammen, d.h. hier sind nur Fellows enthalten, für die Testergebnisse zu allen drei Messzeitpunkten vorliegen.

Im Bereich Unterrichten und Planen wurden vier Testfragen gestellt. Aus den Ergebnissen in Tabelle 33 wird deutlich, dass der Prozentanteil an Personen, die richtig geantwortet haben, für alle Fragen vom ersten zum zweiten Messzeitpunkt zunimmt. Besonders stark fällt dieser Zuwachs für die Fragen TFD1 und TEDS-M4 auf. Dafür gab es bei Frage TEDS-M2 einen leichten Rückgang der richtigen Antwort zum zweiten Messzeitpunkt. Bei dieser, sowie bei TEDS-M4 ist der Anteil an Personen, die zum zweiten Messzeitpunkt richtig geantwortet haben, noch relativ gering. Die Fragen TEDS-M2 und TEDS-M4 sind Fragen, die aus Teilfragen bestehen. Die ausführlichen Ergebnisse zu den Teilfragen finden sich in Tabelle 48 im Anhang. Vom zweiten zum dritten Messzeitpunkt ist für den Bereich Unterrichten – Planen durchgehend eine Verbesserung zu verzeichnen. Diese fällt für TEDS-M2 relativ gering, für TEDS-M4, hingegen, sehr stark aus.



Unterrichten - Planen								
MZP 1 (Welle 1)			MZP 2 (Welle 2)			MZP 3 (Welle 4)		
Frage TFD1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	9	19,57%	richtig	43	93,48%	richtig	43	93,48%
falsch/ nicht gewusst	37	80,43%	falsch/ nicht gewusst	3	6,52%	falsch/ nicht gewusst	3	6,52%
Frage TFD2								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	35	76,09%	richtig	43	93,48%	richtig	44	95,65%
falsch/ nicht gewusst	11	23,91%	falsch/ nicht gewusst	3	6,52%	falsch/ nicht gewusst	2	4,35%
Frage TEDS-M2								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	17	36,96%	richtig	16	34,78%	richtig	18	39,13%
falsch/ nicht gewusst	29	63,04%	falsch/ nicht gewusst	30	65,22%	falsch/ nicht gewusst	28	60,87%
Frage TEDS-M4								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	7	15,22%	richtig	26	56,52%	richtig	37	80,43%
falsch/ nicht gewusst	39	84,78%	falsch/ nicht gewusst	20	43,48%	falsch/ nicht gewusst	9	19,57%

Tabelle 33: Testergebnisse im Bereich Unterrichten – Planen

Für den Bereich Unterrichten – Umsetzen gibt es insgesamt sechs Testfragen, davon drei mit offenem Antwortformat. In diesem Bereich gibt es bei allen Fragen eine Erhöhung des Prozentanteils an Personen, die zum zweiten Messzeitpunkt richtig geantwortet haben. Dies fällt vor allem bei den Fragen TFD12, TFD14, TEDS-M1 und TEDS-M8 auf. Allerdings ist der Anteil an Personen, die richtig geantwortet haben, bei einigen Fragen zum zweiten Messzeitpunkt immer noch sehr klein. Dies trifft vor allem auf die Fragen TEDS-M6, TFD12 und TFD14 zu. Bei den Fragen TFD12, TFD14 und TEDS-M6 handelt es sich um Fragen, die aus mehreren Unterfragen bestehen. Die Häufigkeiten zu den einzelnen Teilfragen werden in Tabelle 49 berichtet. Zum dritten Messzeitpunkt hat sich der Prozentanteil von Personen, die richtig antworten, nur für zwei Fragen erhöht. Bis auf die Frage TEDS-M1 sind die Lösungshäufigkeiten für den Bereich Unterrichten – Umsetzen recht niedrig. Besonders fällt dies für die Fragen TFD12 und TEDS-M6 auf.

Unterrichten - Umsetzen								
MZP 1 (Welle 1)			MZP 2 (Welle 2)			MZP 3 (Welle 4)		
Frage TFD12								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	0	0,00%	richtig	11	23,91%	richtig	6	13,04%
falsch/ nicht gewusst	46	100,00%	falsch/ nicht gewusst	35	76,09%	falsch/ nicht gewusst	40	86,96%
Frage TFD14								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	3	6,52%	richtig	13	28,26%	richtig	26	56,52%
falsch/ nicht gewusst	43	93,48%	falsch/ nicht gewusst	33	71,74%	falsch/ nicht gewusst	20	43,48%
Frage TFD15								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	41	89,13%	richtig	44	97,78%	richtig	-	-
falsch/ nicht gewusst	5	10,87%	falsch/ nicht gewusst	1	2,22%	falsch/ nicht gewusst	-	-
Frage TEDS-M1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	25	54,35%	richtig	37	84,09%	richtig	42	91,30%
falsch/ nicht gewusst	21	45,65%	falsch/ nicht gewusst	7	15,91%	falsch/ nicht gewusst	4	8,70%
Frage TEDS-M6								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	0	0,00%	richtig	2	4,35%	richtig	0	0,00%
falsch/ nicht gewusst	46	100,00%	falsch/ nicht gewusst	44	95,65%	falsch/ nicht gewusst	46	100,0%
Frage TEDS-M8								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	19	41,30%	richtig	34	80,95%	richtig	32	69,57%
falsch/ nicht gewusst	27	58,70%	falsch/ nicht gewusst	8	19,05%	falsch/ nicht gewusst	14	30,43%

Tabelle 34: Testergebnisse im Bereich Unterrichten – Umsetzen

Zum Bereich Leadership wurden zunächst vier Fragen entwickelt. Davon haben zwei ein offenes Antwortformat. Bei allen Fragen nimmt die Zahl der Personen, die richtig geantwortet haben, vom ersten zum zweiten Messzeitpunkt zu. Besonders stark trifft dies auf die Fragen TFD3 und TFD4 zu. Der Prozentanteil der Personen, die TFD4 zum zweiten Messzeitpunkt richtig beantwortet haben, ist allerdings mit 30,77% recht gering. Diese Frage ist eine Frage mit Unterfragen, deren Häufigkeiten in Tabelle 50 im Anhang berichtet werden. Die Lösungshäufigkeiten für den Bereich Leadership nehmen vom dritten zum vierten Messzeitpunkt für zwei Fragen zu und für eine Frage ab. Insgesamt liegen die Lösungshäufigkeiten für diesen Bereich aber relativ niedrig.

Leadership								
MZP 1 (Welle 1)			MZP 2 (Welle 2)			MZP 3 (Welle 4)		
Frage TFD3								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	31	67,93%	richtig	46	100,00%	richtig	-	-
falsch/ nicht gewusst	15	32,61%	falsch/ nicht gewusst	0	0,00%	falsch/ nicht gewusst	-	-
Frage TFD4								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	2	4,35%	richtig	14	30,43%	richtig	5	10,87%
falsch/ nicht gewusst	44	95,65%	falsch/ nicht gewusst	32	69,57%	falsch/ nicht gewusst	41	89,13%
Frage TFD5								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	28	60,87%	richtig	27	58,70%	richtig	32	69,57%
falsch/ nicht gewusst	18	39,13%	falsch/ nicht gewusst	19	41,30%	falsch/ nicht gewusst	14	30,43%
Frage TFD6								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	17	36,96%	richtig	19	41,30%	richtig	21	45,65%
falsch/ nicht gewusst	29	63,04%	falsch/ nicht gewusst	27	58,70%	falsch/ nicht gewusst	25	54,35%

Tabelle 35: Testergebnisse im Bereich Leadership

In der dritten Erhebungswelle wurden zum Bereich Leadership weitere fünf Testfragen erhoben, von denen die erste aus vier Teilfragen besteht, sodass sich insgesamt 8 Testitems

ergeben. Die deskriptiven Statistiken unter Berücksichtigung der Unterfragen finden sich in Tabelle 50 im Anhang.

Leadership - Zusatzfragen								
MZP 1 (Welle 3)			MZP 2 (Welle 4)			MZP 3 (liegt nicht vor)		
Frage TFD16								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	7	16,28%	richtig	11	25,58%	richtig	-	-
falsch/ nicht gewusst	36	83,72%	falsch/ nicht gewusst	32	74,42%	falsch/ nicht gewusst	-	-
Frage TFD17								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	6	13,95%	richtig	6	13,95%	richtig	-	-
falsch/ nicht gewusst	37	86,05%	falsch/ nicht gewusst	37	86,05%	falsch/ nicht gewusst	-	-
Frage TFD18								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	24	55,81%	richtig	27	62,79%	richtig	-	-
falsch/ nicht gewusst	19	44,19%	falsch/ nicht gewusst	16	37,21%	falsch/ nicht gewusst	-	-
Frage TFD19								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	15	34,88%	richtig	16	37,21%	richtig	-	-
falsch/ nicht gewusst	28	65,12%	falsch/ nicht gewusst	27	62,79%	falsch/ nicht gewusst	-	-
Frage TFD20								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	22	51,16%	richtig	19	44,19%	richtig	-	-
falsch/ nicht gewusst	21	48,84%	falsch/ nicht gewusst	24	55,81%	falsch/ nicht gewusst	-	-

Tabelle 36: Testergebnisse im Bereich Leadership - Zusatzfragen

Die fünf neu entwickelten Testitems wurden in der vierten Erhebungswelle in den umfangreichen Test integriert. Für diese Fragen liegen deshalb nur zwei Messzeitpunkte vor. Es ergibt sich auch eine geringere Stichprobengröße, da die Übereinstimmung von Fellows in Welle 3 und Welle 4 nur 43, statt 46 (Übereinstimmung von Fellows in Welle 1, Welle 2 und



Welle 4) beträgt. Mit den Lösungshäufigkeiten für die fünf zusätzlich erhobenen Testitems zum Bereich Leadership verhält es sich ähnlich wie mit den anderen Testfragen zu diesem Bereich. Es gibt zwar hier für vier von fünf Testitems eine Verbesserung zu verzeichnen, aber die Lösungshäufigkeiten zum zweiten Messzeitpunkt fallen insgesamt eher niedrig aus.

Für den Bereich Diagnose gibt es vier Testfragen – alle mit geschlossenem Antwortformat. Zunächst ist festzustellen, dass alle Testfragen mit Ausnahme von TEDS-M7 zum zweiten Messzeitpunkt von mehr Personen richtig beantwortet wurden als zum ersten Messzeitpunkt.

Diagnose								
MZP 1 (Welle 1)			MZP 2 (Welle 2)			MZP 3 (Welle 4)		
Frage TFD10								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	20	43,48%	richtig	44	95,65%	richtig	44	95,65%
falsch/ nicht gewusst	26	56,52%	falsch/ nicht gewusst	2	4,35%	falsch/ nicht gewusst	2	4,35%
Frage TFD11								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	36	78,26%	richtig	41	89,13%	richtig	46	100,0%
falsch/ nicht gewusst	10	21,74%	falsch/ nicht gewusst	5	10,87%	falsch/ nicht gewusst	0	0,00%
Frage TEDS-M5								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	29	63,04%	richtig	40	86,96%	richtig	39	84,78%
falsch/ nicht gewusst	17	36,96%	falsch/ nicht gewusst	6	13,04%	falsch/ nicht gewusst	7	15,22%
Frage TEDS-M7								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	8	17,39%	richtig	8	17,39%	richtig	15	32,61%
falsch/ nicht gewusst	38	82,61%	falsch/ nicht gewusst	38	82,61%	falsch/ nicht gewusst	31	67,39%

Tabelle 37: Testergebnisse im Bereich Diagnose

Die Höhe der Differenz in den Prozentpunkten vom ersten zum zweiten Messzeitpunkt ist jedoch sehr unterschiedlich. Während bei Frage TFD10 52,17 Prozentpunkte zwischen den richtigen Antworten zum ersten und zweiten Messzeitpunkt liegen, sind es bei Frage TFD11 nur 10,87 Prozentpunkte. Bei den Fragen TEDS-M7 und TEDS-M5 handelt es sich um

Testfragen, die aus mehreren Unterfragen zusammengesetzt sind. Die Häufigkeiten für die Unterfragen sind in Tabelle 51 im Anhang berichtet. Vom zweiten zum dritten Messzeitpunkt werden bis auf Frage TEDS-M5, die praktisch gleich gut gelöst wird, alle Fragen häufiger richtig beantwortet als zum zweiten Messzeitpunkt. Mit Ausnahme von Frage TEDS-M7 sind die Lösungshäufigkeiten zum dritten Messzeitpunkt für den Bereich Diagnose insgesamt sehr hoch.

Im Bereich Sprachförderung gibt es zwei Testitems mit geschlossenem Antwortformat. Während bei Frage TFD13 vom ersten zum zweiten Messzeitpunkt eine leichte Erhöhung der mittleren Lösungshäufigkeit zu verzeichnen ist, ändert sich diese bei Frage TEDS-M3 vom ersten zum zweiten Messzeitpunkt nicht. Hier ist allerdings anzumerken, dass Frage TEDS-M3 eine Frage mit mehreren Unterfragen ist. Die genauen Prozentveränderungen der Unterfragen sind in Tabelle 52 im Anhang berichtet. Vom zweiten zum dritten Messzeitpunkt ist bei Frage TFD13 eine leichte Zunahme des Anteils an Personen, die richtig geantwortet festzustellen, für Frage TEDS-M3 eine geringfügige Abnahme. Beide Fragen werden, im Vergleich zu anderen Testfragen, von sehr wenigen Personen richtig beantwortet.

Sprachförderung								
MZP 1 (Welle 1)			MZP 2 (Welle 2)			MZP 3 (Welle 4)		
Frage TFD13								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	1	2,17%	richtig	5	10,87%	richtig	8	17,39%
falsch/ nicht gewusst	45	97,83%	falsch/ nicht gewusst	41	89,13%	falsch/ nicht gewusst	38	82,61%
Frage TEDS-M3								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	10	21,74%	richtig	10	21,74%	richtig	8	17,39%
falsch/ nicht gewusst	36	78,26%	falsch/ nicht gewusst	36	78,26%	falsch/ nicht gewusst	38	82,61%

Tabelle 38: Testergebnisse im Bereich Sprachförderung

Zum Bereich TFD-Einsatz gibt es zwei Testfragen mit geschlossenem Antwortformat<sup>9</sup>. Das prozentuale Verhältnis von richtigen und falschen Antworten hat sich in diesem Bereich bei einer Frage auf sehr hohem Niveau minimal verschlechtert (TFD7) und bei einer Frage deutlich verbessert (TFD9). Frage TFD7 wurde zum dritten Messzeitpunkt nicht mehr erhoben, sodass als Vergleich vom zweiten zum dritten Messzeitpunkt nur noch eine Frage (TFD9) zu betrachten bleibt. Hier haben zum dritten Messzeitpunkt deutlich weniger

<sup>9</sup> Ursprünglich gab es für diesen Themenbereich auch eine offene Testfrage (TFD8), die aber, aufgrund von Inkongruenz zwischen Frage und Antwort-Erwartungshorizont, aus der Bewertung genommen wurde.

Personen richtig geantwortet als zum zweiten Messzeitpunkt, so dass lediglich das Ausgangsniveau des ersten Messzeitpunkts wieder erreicht wird.

TFD Einsatz								
MZP 1 (Welle 1)			MZP 2 (Welle 2)			MZP 3 (Welle 4)		
Frage TFD7								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	43	93,48%	richtig	42	91,30%	richtig	-	-
falsch/ nicht gewusst	3	6,52%	falsch/ nicht gewusst	4	8,70%	falsch/ nicht gewusst	-	-
Frage TFD9								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	30	65,22%	richtig	37	80,43%	richtig	30	65,22%
falsch/ nicht gewusst	16	34,78%	falsch/ nicht gewusst	9	19,57%	falsch/ nicht gewusst	16	34,78%

Tabelle 39: Testergebnisse im Bereich TFD-Einsatz

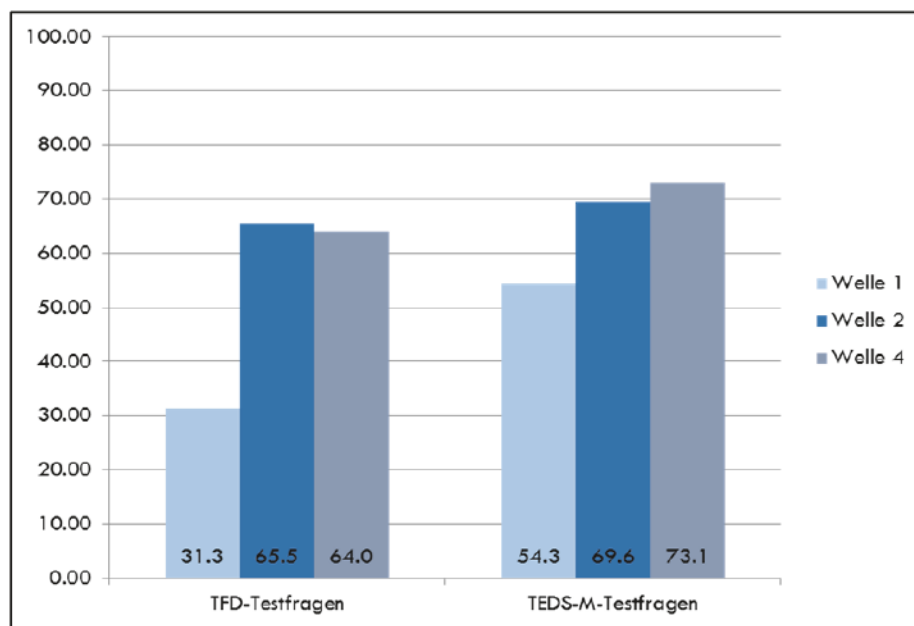


Abbildung 7: Mittlere Lösungshäufigkeit der Testfragen für TFD Fellows über drei Erhebungszeitpunkte

Abbildung 7 zeigt in grafischer Aufbereitung die mittleren Lösungshäufigkeiten für die aufsummierten TEDS-M Fragen sowie für alle TFD Fragen, die zu drei Messzeitpunkten erhoben wurden. Berücksichtigt sind die 46 Fellows, für die zu allen drei Messzeitpunkten Daten vorliegen. Es wird deutlich, dass die Lösungshäufigkeiten vom ersten zum zweiten Messzeitpunkt für die TFD Testfragen sehr viel stärker zunehmen als für die TEDS-M Testfragen. Dies kann teilweise daraus erklärt werden, dass Baseline, also das Ausgangswissen, bei den TEDS-M Testfragen höher liegt. Vom zweiten zum dritten

Messzeitpunkt bleibt die mittlere Lösungshäufigkeit bei den TFD Testfragen praktisch unverändert, während sie für die TEDS-M Fragen noch einmal weiter ansteigt.

Abbildung 9 und Abbildung 10 im Anhang zeigen die mittleren Lösungshäufigkeiten für die Einzelfragen, wobei Fragen, die aus mehreren Unterfragen bestehen, nicht einzeln aufgeschlüsselt werden.

Zusätzlich zu den mittleren Lösungshäufigkeiten können die erreichten Testpunkte der Fellows über die drei Messzeitpunkte betrachtet werden. Bezieht man die Fragen ein, die über alle Wellen erhoben wurden, ergeben sich insgesamt 46 Items. Pro richtig gelöstem Item wurde ein Punkt vergeben, sodass sich die maximal zu erreichende Punktzahl 46 beträgt. Abbildung 8 stellt die erreichten Testpunkte im Wissenstest für jeden Messzeitpunkt als Boxplot dar. Ein Boxplot besteht aus einem Kasten (box) und Zäunen (whiskers), die grafisch die statistische Verteilung einer Variablen darstellen. Die fettgedruckte Linie stellt den Mittelwert der Variable dar. Die Höhe der Box sowie die Zäune sagen etwas über die Streuung der Variable aus. Je höher die Box, bzw. je länger die Zäune, desto stärker variiert oder streut die Variable. Aus Abbildung 8 ist sehr schön abzulesen, dass die Streuung zunächst sehr groß ist, d.h. die Fellows sich in ihrer Testleistung stark unterscheiden. Die Streuung wird mit jedem Messzeitpunkt kleiner, d.h. die Fellows haben zunehmend ähnlichere Testleistungen. Auch der Mittelwert, also die durchschnittliche Testpunktzahl nimmt über die Wellen deutlich zu.

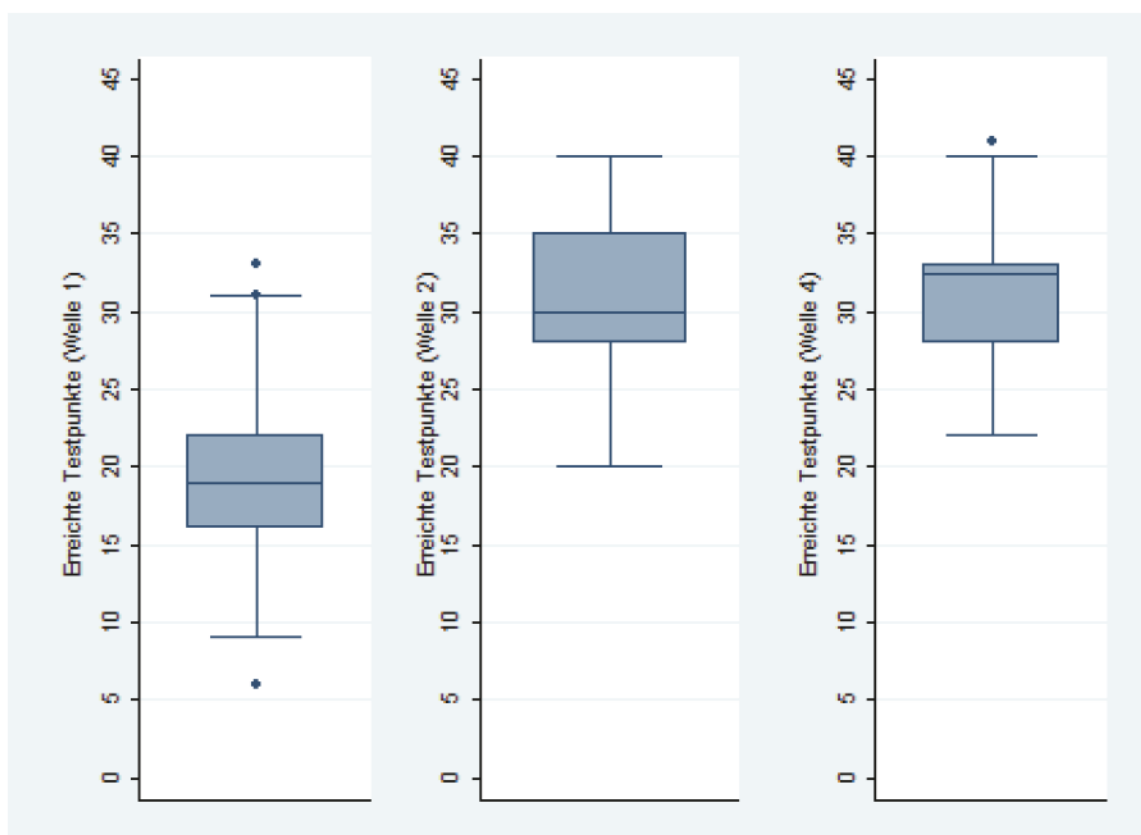


Abbildung 8: Erreichte Testpunkte im Wissenstest über drei Messzeitpunkte



Die genauen deskriptiven Statistiken der Testpunkte finden sich in Tabelle 40. Daraus geht hervor, dass die Fellows im ersten Wissenstest, also in dem Test, der die Lernausgangslage (Baseline) der Fellows messen sollte, im Durchschnitt etwa 20 von 46 Punkten erreichen. Die niedrigste erreichte Punktzahl liegt bei 6 Punkten, die höchste Punktzahl beträgt 33. Die entspricht etwa 71,7%. Zum zweiten Erhebungszeitpunkt ist eine deutliche Verbesserung in der Testleistung zu erkennen. Der Durchschnitt der Testpunkte liegt nun 10 Punkte höher, als bei der Baseline-Messung. Während der Durchschnitt vorher bei 20 Testpunkten lag, stellt dieser Wert in der zweiten Erhebung die niedrigste erreichte Punktzahl dar. Die maximal erreichte Punktzahl beträgt zum zweiten Messzeitpunkt 40 von 46 Punkten. Das entspricht ca. 87,0%. Zum dritten Messzeitpunkt ist eine weitere, wenn auch geringfügige, Verbesserung der Fellows zu erkennen. Die durchschnittlich erreichte Testpunktzahl beträgt 31 Punkte, die niedrigste erreichte Punktzahl liegt bei 22 Punkten, die beste Testleistung beträgt 41 Punkte. Dies entspricht ca. 89,1%.

Testpunkte								
MZP 1 (Welle 1)			MZP 2 (Welle 2)			MZP 3 (Welle 4)		
M	Min	Max	M	Min	Max	M	Min	Max
19,85	6	33	30,83	20	40	31,00	22	40

Tabelle 40: Testpunkte - Deskriptive Statistiken

## 8.4 Professionalität von TFD Fellows im Vergleich

Um die Ergebnisse der Fellows mit Ergebnissen anderer Studien vergleichen zu können, wurden in den Wissenstest 8 Testfragen eines bereits mehrfach erprobten Tests integriert. Es handelt sich dabei um den Test zum Pädagogischen Unterrichtswissen (PUW) aus der TEDS-M Studie. Des Weiteren sollten die Fellows mit einer unabhängigen Stichprobe verglichen werden. Dafür hat das Forscherteam eine Gruppe von Lehramtsstudierenden befragt, die Stipendiaten der Stiftung der Deutschen Wirtschaft (SDW) sind. Der Fragebogen für diese Gruppe beinhaltete den ausführlichen Wissenstest mit TEDS-M Fragen sowie eine Auswahl<sup>10</sup> von TFD Testfragen.

### 8.4.1 TFD Fellows im Vergleich: Die TEDS-M Studie

Die TEDS-M Studie (*Teacher Education and Development Study – Learning to Teach Mathematics*) ist eine internationale Vergleichsstudie im Auftrag der IEA (International Association for the Evaluation of Education), die das mathematische Fachwissen, mathematikdidaktische Wissen sowie *Pädagogisches Unterrichtswissen* angehender Mathematiklehrkräfte der Sekundarstufe und angehender Primarstufenlehrkräfte am Ende ihrer Ausbildung erfasst<sup>11</sup>. Für die deutsche Stichprobe liegt dieser Zeitpunkt kurz vor dem 2. Staatsexamen (2008). Des Weiteren wurde der PUW-Test in einer Ergänzungsstudie von TEDS-M eingesetzt: der LEK Studie (*Längsschnittliche Entwicklung pädagogischer*

<sup>10</sup> Einige Fragen beziehen sich auf sehr spezifische TFD Inhalte, bzw. TFD-Materialien und sind nur für Fellows zu beantworten.

<sup>11</sup> Teilgenommen haben insgesamt 18 Länder: Botswana, Kanada, Chile, Taiwan, Georgien, Deutschland, Malaysia, Norwegen, Oman, die Philippinen, Polen, Russland, Singapur, Spanien, die Schweiz (nur die deutschsprachigen Kantone), Thailand, und die Vereinigten Staaten von Amerika.

*Kompetenzen von Lehramtsstudierenden*). Hier wird das *Pädagogische Unterrichtswissen von* Lehramtsstudierenden aller Fächer im ersten Semester (Ende 2008/Anfang 2009) getestet. Uns liegen zum Vergleich mit den TFD-Fellows die Testergebnisse dieser drei Gruppen vor. Tabelle 41 zeigt zunächst den Vergleich der Lernausgangslage, d.h. hier wird die Stichprobe der Lehramtsstudierenden im 1. Semester mit den TFD Fellows zum ersten Messzeitpunkt (vor dem Online Campus) verglichen. Die Ergebnisse für beide Gruppen stellen gleichsam die jeweilige Baseline des Pädagogischen Unterrichtswissens dar. Die Items, bei denen die Gruppe der Fellows besser abschneidet als die LEK-Gruppe, sind fett gedruckt.

Item	Lösungshäufigkeit in % für die betreffende Stichprobe	
	LEK (Lernausgangslage, 1. Semester)	TFD (MZP 1) (Lernausgangslage, vor Online Campus)
TEDS-M1	36,87	50,77
TEDS-M2.1	50,86	93,85
TEDS-M2.2	47,16	93,85
TEDS-M2.3	54,57	95,38
TEDS-M2.4	43,70	83,08
TEDS-M2.5	23,46	47,69
TEDS-M3.1	64,92	84,62
TEDS-M3.1.1	65,71	71,88
TEDS-M3.2	37,96	38,98
TEDS-M3.2.1	31,94	70,00
TEDS-M4.1	25,89	29,23
TEDS-M4.2	24,87	30,77
TEDS-M4.3	29,44	29,23
TEDS-M4.4	29,19	18,46
TEDS-M5.1	57,21	70,77
TEDS-M5.2	75,55	92,31
TEDS-M6.1	1,01	1,54
TEDS-M7.1	43,43	61,54
TEDS-M7.2	44,19	61,54
TEDS-M7.3	38,99	44,62
TEDS-M7.4	20,30	35,38
TEDS-M8	37,21	40,00
N <sub>Max</sub>	415	65

Tabelle 41: Vergleich der Testergebnisse zum Pädagogischen Unterrichtswissen von Lehramtsstudierenden und TFD Fellows – Lernausgangslage

Dadurch fällt auf den ersten Blick auf, dass die Fellows bei fast allen Fragen besser abschneiden als die Gruppe der Lehramtsstudierenden. Und das, obwohl in der LEK Studie

darauf hingewiesen wird, dass die Ergebnisse der Lehramtsstudierenden vermutlich eher besser sind, als in der Grundgesamtheit der Lehramtsstudierenden, da es sich um eine Gelegenheitsstichprobe handelt. Die Studierenden wurden also nicht zufällig ausgewählt, sodass die Autoren davon ausgehen, dass es sich eher um besonders motivierte Studierende handelt, da sie sich von selbst zur Testteilnahme gemeldet haben.

Tabelle 42 zeigt den Vergleich am Ende der Ausbildung zwischen den beiden deutschen TEDS-M Stichproben (Primarstufe und Sekundarstufe) und den TFD Fellows zu zwei Messzeitpunkten: direkt nach dem Online Campus und vier Monate nach Ende der Sommerakademie. Die Items, bei denen die Gruppe der Fellows zu beiden Messzeitpunkten besser abschneidet als die anderen Gruppen, sind fettgedruckt.

Item	Lösungshäufigkeit in % für die betreffende Stichprobe			
	TEDS-M (Primar) (Leistung kurz vor 2. Staatsexamen)	TEDS-M (Sekundar) (Leistung kurz vor 2. Staatsexamen)	TFD (MZP 2) (Leistung nach Online Campus)	TFD (MZP 3) (Leistung 4 Mon. nach Sommerakademie)
TEDS-M1	66,29	57,88	83,87	<b>91,67</b>
TEDS-M2.1	93,73	87,14	95,38	<b>100,00</b>
TEDS-M2.2	85,22	88,42	92,31	<b>97,29</b>
TEDS-M2.3	96,76	93,85	98,46	<b>100,00</b>
TEDS-M2.4	80,27	84,05	86,15	<b>85,42</b>
TEDS-M2.5	39,48	46,48	46,15	45,83 S
TEDS-M3.1	69,70	64,79	86,15	<b>85,11</b>
TEDS-M3.1.1	63,38	59,29	81,25	<b>67,44</b>
TEDS-M3.2	47,73	67,05	42,86	51,11 S
TEDS-M3.2.1	50,03	56,51	63,49	<b>65,91</b>
TEDS-M4.1	84,64	79,06	87,69	<b>95,83</b>
TEDS-M4.2	92,97	84,68	84,62	89,58 P
TEDS-M4.3	95,14	89,50	84,62	93,75 P
TEDS-M4.4	83,35	74,25	52,31	81,25 P
TEDS-M5.1	79,08	78,06	89,23	<b>89,58</b>
TEDS-M5.2	96,19	90,67	95,38	91,67 P
TEDS-M6.1	29,49	24,57	4,62	0,00 P/S
TEDS-M7.1	89,37	83,39	73,44	77,08 P/S
TEDS-M7.2	85,58	82,36	76,56	83,33 P
TEDS-M7.3	80,63	77,84	67,19	75,00 P/S
TEDS-M7.4	49,97	42,04	37,50	<b>56,25</b>
TEDS-M8	40,43	53,24	72,13	<b>66,67</b>
N <sub>Max</sub>	3718	2292	63	46

**Tabelle 42: Vergleich der Testergebnisse zum Pädagogischen Unterrichtswissen von Referendaren und TFD Fellows - Leistung am Ende der Ausbildung**

Wenn die Fellows nicht durchgehend besser abschneiden, wird die TEDS-M (Teil-)gruppe spezifiziert, gegenüber der die Fellows schlechter abschneiden. Dafür wird hinter der Lösungshäufigkeit für die Fellows der Buchstabe der TEDS-M Stichprobe gesetzt: S = Stichprobe der Sekundarschullehrkräfte hat bei diesem Item besser abgeschnitten; P = Stichprobe der Primarschullehrkräfte hat bei diesem Item besser abgeschnitten. Insgesamt erzielen die TFD Fellows bei etwa drei Viertel der Testfragen höhere durchschnittliche Lösungshäufigkeiten als Referendare der Sekundarstufe und bei etwa zwei Drittel der Testfragen erzielen die TFD Fellows höhere durchschnittliche Lösungshäufigkeiten als Referendare der Primarstufe.

#### **8.4.2 TFD-Fellows im Vergleich mit Stipendiaten der Stiftung der Deutschen Wirtschaft**

Im Folgenden werden die Ergebnisse der Fellows mit den Ergebnissen einer unabhängigen Stichprobe verglichen werden. Dabei handelt es sich um Lehramtsstudierende der Ludwig-Maximilian Universität München, die Stipendiaten der Stiftung der Deutschen Wirtschaft (SDW) sind. Die Gruppe wurde mittels Online-Instrument befragt. Insgesamt haben 37 Personen an der Befragung teilgenommen. Die meisten der Befragten sind weiblich (73,0%) und der Altersdurchschnitt der Stichprobe liegt bei etwa 24 Jahren. Die meisten der Teilnehmenden befinden sich bereits in höheren Semestern – 25 Personen studieren im sechsten Semester oder höher. Insgesamt reicht die Spannweite der Semester vom ersten bis zum elften Semester.

Die Tatsache, dass die Teilnehmenden nicht alle am Anfang oder am Ende ihres Studiums stehen, relativiert die Aussagekraft des Vergleichs. Die Gruppe noch einmal nach der Semesterzahl aufzuteilen, ist jedoch bei der geringen Stichprobengröße nicht methodisch nicht sinnvoll.

Die Kollegiaten der SDW bilden eine Vergleichsgruppe aus Lehramtsstudierenden, die sich von den 'durchschnittlichen' Studierenden dadurch unterscheiden, dass sie aufgrund ihrer Studienleistungen und ihres ehrenamtlichen Engagements für ein Stipendium ausgewählt wurden. Daher können aus den Vergleichen der Wissenstests Anhaltspunkte für die Qualität der Auswahlverfahren in beiden Gruppen abgeleitet werden. Voraussetzung dafür ist allerdings, dass die Gruppen von ihrem Studienhintergrund vergleichbar sind. Da dies vor allem aufgrund der Freiwilligkeit der Teilnahme nicht sichergestellt werden konnte, müssen die im Folgenden präsentierten Ergebnisse vorsichtig interpretiert werden.

Bei den TEDS-M Fragen, liegen die Lösungshäufigkeiten der Fellows für 8 von 22 Fragen höher als die der SDW-Stichprobe. Dies wird am deutlichsten bei TEDS-M1 und TEDS-M5.1. Bei 7 Fragen unterscheiden sich die beiden Stichproben praktisch nicht und bei den übrigen 7 Fragen erzielt die SDW-Stichprobe höhere Lösungshäufigkeiten.



Item	Lösungshäufigkeit in % für die betreffende Stichprobe			
	TFD (MZP 1) (Lernausgangslage, vor Online Campus)	TFD (MZP 2) (Leistung nach Online Campus)	TFD (MZP 3) (Leistung 4 Mon. nach Sommerakademie)	SDW (MZP 1)
TEDS-M1	50,77	83,87	91,67	72,97
TEDS-M2.1	93,85	95,38	100,00	100,00
TEDS-M2.2	93,85	92,31	97,29	97,30
TEDS-M2.3	95,38	98,46	100,00	100,00
TEDS-M2.4	83,08	86,15	85,42	94,59
TEDS-M2.5	47,69	46,15	45,83	43,24
TEDS-M3.1	84,62	86,15	85,11	86,49
TEDS-M3.1.1	71,88	81,25	67,44	62,16
TEDS-M3.2	38,98	42,86	51,11	77,78
TEDS-M3.2.1	70,00	63,49	65,91	75,00
TEDS-M4.1	29,23	87,69	95,83	86,49
TEDS-M4.2	30,77	84,62	89,58	94,59
TEDS-M4.3	29,23	84,62	93,75	86,49
TEDS-M4.4	18,46	52,31	81,25	86,49
TEDS-M5.1	70,77	89,23	89,58	70,27
TEDS-M5.2	92,31	95,38	91,67	91,89
TEDS-M6.1	1,54	4,62	0,00	13,51
TEDS-M7.1	61,54	73,44	77,08	78,38
TEDS-M7.2	61,54	76,56	83,33	83,78
TEDS-M7.3	44,62	67,19	75,00	83,78
TEDS-M7.4	35,38	37,50	56,25	48,65
TEDS-M8	40,00	72,13	66,67	59,46
N	65	63	46	37

Tabelle 43: Vergleich der Testergebnisse zum Pädagogischen Unterrichtswissen von TFD Fellows und SDW Stipendiaten

Für die SDW-Stichprobe liegen auch Ergebnisse zu den Fragen vor, die basierend auf den TFD Materialien entwickelt wurden. Hier erreichen die Fellows fast durchgehend bessere Lösungshäufigkeiten. Am deutlichsten ist dies bei Frage TFD1 und bei Frage TFD14 zu erkennen. Durch diese Analyse wird die curriculare Validität der selbst Entwickelten Testfragen zu TFD Materialien bestätigt.

Item	Lösungshäufigkeit in % für die betreffende Stichprobe			
	TFD (MZP 1) (Lernausgangslage, vor Online Campus)	TFD (MZP 2) (Leistung nach Online Campus)	TFD (MZP 3) (Leistung 4 Mon. nach Sommerakademie)	SDW (MZP 1)
TFD1	16,92	86,15	<b>93,48</b>	54,05
TFD2	70,77	89,23	<b>95,65</b>	91,89
TFD5	64,62	67,69	<b>69,57</b>	64,68
TFD9	67,69	81,25	65,22	81,08
TFD10	40,00	98,41	<b>95,65</b>	75,68
TFD11	75,38	93,55	<b>100</b>	97,30
TFD12	0,00	23,08	<b>13,04</b>	2,70
TFD14	1,54	21,54	<b>56,52</b>	0,00
<b>N</b>	<b>65</b>	<b>63</b>	<b>46</b>	<b>37</b>

Tabelle 44: Vergleich der Testergebnisse der TFD Wissensfragen von TFD Fellows und SDW Stipendiaten

## 9 Fazit

Die Maßnahmen zur Qualifizierung von Teach First Fellows weisen insgesamt eine hohe Akzeptanz auf. Dies gilt sowohl für die gewählten Formate (Online-Campus, Präsenzphase Sommerschule, Feriencamp für Schüler MSact!, Austausch zwischen Fellows, Mentoring und Weiterbildung während des Einsatz) als auch für die Inhalte wie für die Interaktion im Prozess der Weiterbildung. Die mittleren Bewertungen zu entsprechenden Aussagen in der Teilnehmerbefragung liegen weit überwiegend im deutlich positiven Bereich.

Die Möglichkeit zur offenen Kommentierung wurde von einigen Fellows genutzt um bestehendes Optimierungspotential anzumerken. Dies wird erstens in Bezug auf die theorie- und wissenschaftsbezogene Aufbereitung einzelner Inhaltfelder gesehen, zweitens in Bezug auf die technische Umsetzung der Online-Angebote und drittens in Bezug auf die Anleitung von konkreter Praxis.

Die Herausforderung für die weitere theorie- und wissenschaftsbezogene Optimierung der Materialien besteht darin, einer Gruppe von Universitätsabsolvent/inn/en mit heterogenen Vorerfahrungen in Bezug auf das Tätigkeitsfeld ein Angebot zu unterbreiten, das die in den jeweiligen Studienfächern ausgebildeten Aneignungsstrategien für neue Inhalte unterstützt. Hier gilt es die Balance zwischen einerseits eingeforderter Differenzierung und Vertiefung und andererseits handlungsorientierender Klarheit zu halten.

Die Herausforderung für die weitere Entwicklung der Online-Angebote für Fellows ist (vorausgesetzt die Ressourcen zur Programmierung und Pflege sind gegeben) wohl am leichtesten umzusetzen. Auch hier ist jedoch darauf zu achten, dass die konkreten Verbesserungswünsche jeweils nur von einzelnen formuliert wurden und keinen Rückschluss auf mehrheitlich vorhandenen Nutzungsgewohnheiten und -bereitschaften der Online-Angebote zum Ausdruck bringen.

Die Herausforderung in Bezug auf eine konkretere Praxisanleitung besteht darin, dass die Einsatzprofile der Fellows sehr vielfältig sind und sich nur partiell überschneiden. Abgesehen davon, ob standardisierte Interventionspläne generell als breitenwirksam erachtet werden oder nicht, besteht hier die Problematik der Passung von im Einzelfall erwünschter Standardisierung und variablen Einsatzplänen. Deshalb könnten gegenwärtig nur für Teilbereiche des Fellow-Einsatzes stärkere Standardisierungen erprobt werden.

Jenseits der Akzeptanz durch die Teilnehmenden an der Qualifizierung fokussiert die vorliegende Evaluationsstudie auf die summativen Ergebnisse hinsichtlich Wissensaufbau, Kompetenzerleben und Selbstwirksamkeitserwartungen. Aus einer psychometrischen Perspektive stellt dabei das getestete Wissen den härtesten Indikator für Erfolg dar. Der Testwert ist weniger als selbst eingeschätzte Kompetenzen und Wirkungen durch mögliche Beurteilungsfehler verzerrt. Gleichwohl sind die Selbsteinschätzungen wichtige Ergänzungen

zum Test, weil sie die subjektive Perspektive einfangen und über die kognitiven Effekte bei den Fellows hinausweisen auf deren künftiges Verhalten und Selbstbewusstsein.

Die Teach First Sommerakademie erweist sich als eindeutig funktional im Hinblick auf den intendierten Wissensaufbau. Die Kenntnisse der Fellows zu den Inhalten der Sommerakademie haben sich schon im Verlauf der Online-Beschäftigung mit den bereitgestellten Materialien verdoppelt. Der Wissensaufbau ist auch mehrere Monate nach der ursprünglichen Aneignungsphase noch stabil. Dieses Ergebnis ist umso erfreulicher, als ein paralleler Test mit Fragen aus einer großen Repräsentativbefragung von Lehramtskandidat/inn/en zeigt, dass die TFD-Fellows schon zu Beginn ihrer Qualifizierung ein überdurchschnittliches Wissen zu Schule und Unterricht mitbringen, und ihr pädagogisches Unterrichtswissen im gestellten Testausschnitt nach einem halben Jahr dem von Lehrkräften im Vorbereitungsdienst entspricht.

Ein leichter Rückgang der erwarteten Selbstwirksamkeit bei den Fellows kann beobachtet werden, wenn die Werte vor der Qualifizierung und nach den ersten Monaten im Einsatz verglichen werden. Auch wenn sich die Fellows zu allen Messzeitpunkten und in allen erfragten Tätigkeitsbereichen als eher wirksam denn unwirksam empfinden, so ist doch insbesondere in Bezug auf Schülermotivation und Klassenführung der Ausblick auf die eigene Wirksamkeit nach einem halben Jahr etwas reduziert. Dieser Befund zeigt einen scheinbaren Kontrast zum gelingenden Wissensaufbau und steht vermutlich mit der vertieften Erfahrung und Kenntnis von real vorfindlichen Bedingungen schulischer Arbeit in Zusammenhang. Diese Entwicklung der Selbstwirksamkeit ist für den ersten längeren Praxiskontakt mit schulischen Handlungsfeldern nicht ungewöhnlich und kann als solche auch in das Qualifizierungsprogramm von Teach First einbezogen werden.

Positiv entwickelt sich hingegen im Verlauf der Qualifizierung die wahrgenommene Unterrichtskompetenz der Fellows. D.h. die Fellows sehen sich im Verlauf mehr und mehr in der Lage alltägliche Anforderungen an die Unterrichtstätigkeit gut zu erfüllen. Eventuell gibt es in diesem Bereich aber noch ein nicht ausgeschöpftes Entwicklungspotential. Die Formulierung „eventuell“ ist dabei bewusst vorsichtig gewählt, weil zwar Vergleichswerte aus größeren Studien existieren, die die Annahme eines solchen Entwicklungspotentials nahe legen, hier jedoch berücksichtigt werden muss, dass die Fellows an Schulen in schwieriger Lage arbeiten, in denen beginnende Lehrkräfte vor dem Hintergrund der realen Probleme vielleicht zu einer geringeren selbsteingeschätzten Kompetenz neigen als im Mittel aller Schulen.



## 10 Ausblick

Am Ende dieses deskriptiven Berichts stellt sich die Frage, welche weiteren Fragen erkennbar wurden und welche Analysen zur weiteren Optimierung von Qualifizierungsangeboten für pädagogische Tätigkeiten in der Schule hilfreich sein können. Diese sollen hier nur ansatzweise und als Stichpunkte aufgeführt werden:

1. Analysen zur Weiterentwicklung von Online-Materialien, z.B. Themen: Persönlichkeit, Leistungsdiagnostik, Leadership
  - ☐ Inwiefern erreichen überarbeitete Materialien in den angezeigten Bereichen eine höhere Akzeptanz und tragen zu einer erhöhten Selbstwirksamkeit bei?
  - ☐ Wie wird die spezifische Herausforderung von Schulen in schwieriger Lage adressiert (z.B. Ansätze der Sprachförderung, Gründe für Leistungsdefizite bei niedrigem SES und entsprechende Interventionen)
  - ☐ Wie entwickelt sich das Verständnis von Bildungsbenachteiligung bei den Fellows (Hutchings et al.2006)
2. Analysen zum Zusammenhang von schulischer Qualifikation der Fellows und der Entwicklung von beruflichen Perspektiven während und im Anschluss an den Einsatz
  - ☐ Produktive Entwicklung der spezifischen Rolle von Fellows im Schulalltag (z.B. Radfort et al. 2015)
  - ☐ Verbleib im Lehrerberuf oder in anderen pädagogischen Tätigkeiten; Engagement für Bildungsqualität nach dem Durchlaufen des Programms (Hutchings et al 2006)
3. Analysen zur Bedeutung von Einsatzprofilen für den Qualifizierungserfolg
  - ☐ Inwiefern lassen sich aus unterschiedlichen Einsatzprofilen Vorhersagen für den Erfolg der Qualifizierung hinsichtlich Wissen, Unterrichtskompetenz und Selbstwirksamkeit ableiten? Im Hintergrund steht hier die Überlegung, dass sich die Qualifizierung nicht nur durch formal gestaltete Angebote vollzieht, sondern auch durch die Art der Einbindung in die Handlungskontexte (Schicke 2012).
4. Analysen zur Bedeutung von Selbstwirksamkeit im Qualifizierungsprozess
  - ☐ Wie lässt sich die Entwicklung der Selbstwirksamkeit erklären? Stress und Arbeitsbelastung als mögliche Prädiktoren für Selbstwirksamkeit (Klassen, Durksen, 2014)
  - ☐ Inwiefern ist Selbstwirksamkeit für Kompetenzerwerb bedeutsam? Zur Vorhersage von Kompetenz und Wissen oder als Effekt aus Kompetenz und Wissen? (Klassen, Tse, 2014; Holzberger, Phillipp, Kunter, 2013)

- ☐ Dient Selbstwirksamkeit zur Vorhersage für Verpflichtung gg. TFD? Welle 2 als Prädiktor für die Teilnahme an Welle 3 (Chesnut, Burley, 2015)

5. Vertiefende Analysen zu weiter reichenden Effekten von TFD

- ☐ Analyse von wahrgenommener Unterrichtsqualität durch SchülerInnen (z.B. Holzberger, Phillipp, Kunter, 2013). Wenn eine standardisierte Intervention implementiert würde, könnten auch Schülerleistungen sinnvoll in die Evaluation einbezogen werden.
- ☐ Analyse von übertragbaren Elementen auf die zweite Phase der Lehrerbildung (Hutchings et al. 2006)

## 11 Literaturverzeichnis

Abs, H. J., Anderson-Park, E. (2014). Programme zur Berufseinführung von Lehrpersonen. In: Terhart, Ewald; Bennewitz Hedda; Rothland, Martin (Hrsg.): *Handbuch der Forschung zum Lehrberuf. 2. Auflage*. Münster (Waxmann).

Abs, H.J., Döbrich, P., Gerlach-Jahn, A., & Klieme, E. (2009). *Pädagogische Entwicklungsbilanzen an Studienseminaren. Auswahl und statistische Analyse der Erhebungsinstrumente* (Materialien zur Bildungsforschung, Bd. 22). Frankfurt: Deutsches Institut für Internationale Pädagogische Forschung.

Chesnut, S. R. & Burley, H. (2015). Self-efficacy as a predictor of commitment to the teaching profession: A meta-analysis. *Educational Research Review*, 15, 1–16.

Fend, H. (2008). *Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen*, Wiesbaden: VS Verlag für Sozialwissenschaften.

Hutchings, M., Maylor, U., Mendick, H., Menter, I. & Smart, S. (2006). *An evaluation of innovative approaches to teacher training on the Teach First programme: Final report to the Training and Development Agency for Schools*, London: Institute for Policy Studies in Education.

Keller, M. M., Goetz, T., Becker, E.S.; Morger, V. & Hensley, L. (2014). Feeling and showing: A new conceptualization of dispositional teacher enthusiasm and its relation to students' interest, *Learning and Instruction*, 33, 29–38.

Klassen, R. M. & Durksen, T. L. (2014). Weekly self-efficacy and work stress during the teaching practicum: A mixed methods study. *Learning and Instruction*, 33, 158–169.

Klassen, R. M. & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, 12, 59–76.

König, J. (2014). Forschung zum Erwerb von pädagogischem Wissen angehender Lehrkräfte in der Lehrerbildung. In: Terhart, E.; Bennewitz, H.; Rothland, M. (Hg.). *Handbuch der Forschung zum Lehrberuf*. 2. Aufl. Münster: Waxmann. 615-641.

König, J., Blömeke, S. & Kaiser, G. (2015). Early Career Mathematics Teachers' General Pedagogical Knowledge and Skills: Do Teacher Education, Teaching Experience, and Working Conditions Make a Difference? *International Journal of Science and Mathematics Education*, 13 (2), 331–350.

König, J. & Blömeke, S. (2010). *Pädagogisches Unterrichtswissen (PUW). Dokumentation der Kurzfassung des TEDS-M Testinstruments zur Kompetenzmessung in der ersten Phase der Lehrerbildung*. Berlin: Humboldt-Universität zu Berlin.

Krauss, S., Neubrand, M., Blum, W., Baumert, J., Brunner, M., Kunter, M. et al. (2008). Die Untersuchung des professionellen Wissens deutscher Mathematik-Lehrerinnen und -Lehrer im Rahmen der COACTIV-Studie. *Journal für Mathematikdidaktik*, 29(3/4), 223–258.

McConney, A., Price, A. & Woods-McConney, A. (2012). *Fast Track Teacher Education. A Review of the Research Literature on Teach for All Schemes*. New Zealand Post-Primary Teacher's Association (PPTA), Perth: Centre for Learning, Change and Development.

Neuweg, H. G. (2014). Das Wissen der Wissenvermittler. Problemstellung, Befunde und Perspektiven der Forschung zum Lehrwissen. In: Terhart, E.; Bennewitz, H.; Rothland, M. (Hg.). *Handbuch der Forschung zum Lehrerberuf*. 2. Aufl. Münster: Waxmann. 583-614.

Radford, J., Bosanquet, P., Webster, R. & Blatchford, P. (2015). Scaffolding learning for independence: Clarifying teacher and teaching assistant roles for children with special educational needs. *Learning and Instruction*, 36, 1–10.

Rolloff Henoch, J., Klusmann, U., Lüdtke, O. & Trautwein, U. (2015). Who becomes a teacher? Challenging the “negative selection” hypothesis. *Learning and Instruction*, 36, 46–56.

Schicke, H. (2012). *Organisation als Kontext der Professionalität: Beruflichkeit pädagogischer Arbeit in der Transformationsgesellschaft*. Bielefeld: W. Bertelsmann.

Tschannen-Moran, M. & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, 17, 783-805.



## 12 Anhang

<b>Persönliche Voraussetzungen</b>					
<b>Geschlecht</b>	<b>Häufigkeiten</b>				
	<b>absolut</b>	<b>relativ</b>	<b>kumuliert</b>		
weiblich	39	60,00%	60,00%		
männlich	26	40,00%	100,00%		
	65	100,00%			
<b>Muttersprache</b>	<b>Häufigkeiten</b>				
	<b>absolut</b>	<b>relativ</b>	<b>kumuliert</b>		
ausschließlich Deutsch	52	80,00%	80,00%		
Deutsch und eine andere Sprache	11	16,92%	96,92%		
ausschließlich eine andere Sprache	2	3,08%	100,00%		
	65	100,00%			
<b>Vorbereitungsdienst für Lehrämter</b>	<b>Häufigkeiten</b>				
	<b>absolut</b>	<b>relativ</b>	<b>kumuliert</b>		
Teilnahme: ja	0	0,00%	0,00%		
Teilnahme: nein	65	100,00%	100,00%		
	65	100,00%			
<b>Prüfungsverfahren für Lehrämter</b>	<b>Häufigkeiten</b>				
	<b>absolut</b>	<b>relativ</b>	<b>kumuliert</b>		
Teilnahme: ja	0	0,00%	0,00%		
Teilnahme: nein	65	100,00%	100,00%		
	65	100,00%			
	<b>N</b>	<b>Mean</b>	<b>Stdv</b>	<b>Min</b>	<b>Max</b>
<b>Alter</b>	65	27,77	3,09	24	38
<b>Abiturnote</b>	65	1,94	0,63	1,00	3,40
<b>Abschlussnote erstes Fach</b>	65	1,55	0,43	1,00	3,10
<b>Abschlussnote zweites Fach</b>	38	1,46	0,38	1,00	3,00
<b>Arbeit mit Kindern im schulischen Kontext (in Monaten)</b>	39	10,31	12,73	1	72
<b>Arbeit mit Kindern im außerschulischen Kontext (in Monaten)</b>	53	31,43	30,72	1	130
<b>Auslandsaufenthalt zum Studieren oder Arbeiten (in Monaten)</b>	56	22,21	19,25	1	80
<b>Merkmale der Nutzung der Lernumgebung</b>					
<b>Häufigkeit Kontakt zu anderen Fellows über Lernplattform</b>	<b>Häufigkeiten</b>				
	<b>absolut</b>	<b>relativ</b>	<b>kumuliert</b>		
mehrmals am Tag	1	1.56%	1.56%		
einmal am Tag	5	7.81%	9.38%		
mehrmals pro Woche	30	46.88%	56.25%		
einmal pro Woche	15	23.44%	79.69%		
weniger als einmal pro Woche	13	20.31%	100%		
	64	100,00%			

Tabelle 45: Deskriptive Statistiken

Sprach- und Kulturwissenschaften	Rechts-, Wirtschafts- und Sozialwissenschaften	Mathematik, Naturwissenschaften	Ingenieurwissenschaften
<ul style="list-style-type: none"> <li>• Angewandte Kulturwissenschaften</li> <li>• BA Sozial- und Kulturanthropologie</li> <li>• Bachelor Germanistik</li> <li>• Angewandte Medienwissenschaften</li> <li>• Bachelor of Liberal Arts</li> <li>• Diplom Psychologie</li> <li>• Bachelor of Life Science</li> <li>• Bachelor Philosophie</li> <li>• Diplom Katholische Theologie</li> <li>• European Studies</li> <li>• Europäische Ethnologie</li> <li>• Slawistik</li> <li>• Islamwissenschaft</li> <li>• Kommunikationsdesign/Zeitbasierte Medien</li> <li>• Latein</li> <li>• Spanische Studien</li> <li>• Vergleichende Kultur- und Religionswissenschaft</li> <li>• Bildungswissenschaft</li> <li>• Interkulturelle Kommunikation</li> <li>• Lateinamerikastudien</li> <li>• Mehrsprachigkeit und Bildung</li> <li>• Postcolonial Literary and Cultural Studies</li> <li>• Sozial- und Kulturanthropologie</li> <li>• Deutsche Literatur</li> </ul>	<ul style="list-style-type: none"> <li>• BA Regionalstudien Asien/Afrika</li> <li>• Sozialwissenschaften</li> <li>• Staatswissenschaften</li> <li>• BA Soziologie</li> <li>• Politikwissenschaften</li> <li>• BA Europa-Studien</li> <li>• BA Business Administration</li> <li>• BA of Public Administration</li> <li>• Betriebswirtschaftslehre</li> <li>• International Economics</li> <li>• International Management</li> <li>• Internationale Beziehungen</li> <li>• Intercultural Conflict Management</li> <li>• Politik- und Verwaltungswissenschaften</li> <li>• Public Economics, Law and Politics</li> <li>• Politik, Integration, Sicherheit</li> <li>• Environmental and Natural Resource Economics</li> <li>• International and European Studies</li> <li>• Nordeuropastudien</li> <li>• Rechtswissenschaften</li> <li>• Sozialwissenschaften</li> <li>• Volkswirtschaftslehre</li> <li>• Wirtschaftsingenieurwesen</li> </ul>	<ul style="list-style-type: none"> <li>• BA of Science</li> <li>• Geographie</li> <li>• Geologie</li> <li>• Mathematik</li> </ul>	<p>Maschinenbau</p>

Tabelle 46: Kategorisierung der Studienabschlüsse der TFD Fellows

	Nur in Welle 1	Nur in Welle 3	Nur in Welle 4	Nur in Welle 1 und 2	Nur in Welle 1 und 3	Nur in Welle 2 und 3	Nur in Welle 2 und 4	Nur in Welle 3 und 4	Nur in Welle 1, 2 und 4
ID	41	Neu3, Neu5, 0	Unbekannt1, Unbekannt2, Unbekannt3, Unbekannt4, Unbekannt5, Unbekannt6, Unbekannt7, Unbekannt8	20, 30, 73	44	Neu2	Neu1	Neu4	3, 10, 17,21

Tabelle 47:ID-Nummern der Personen, die nicht in allen Befragungswellen enthalten sind

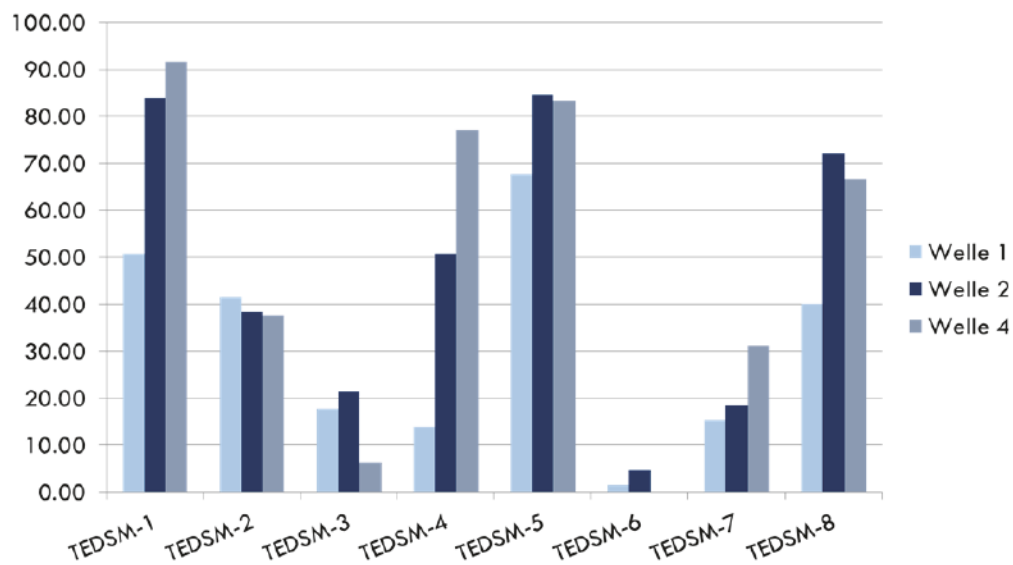


Abbildung 9: Wissensfragen aus PUW-Test der TEDS-M Studie - Lösungshäufigkeit über drei Messzeitpunkte

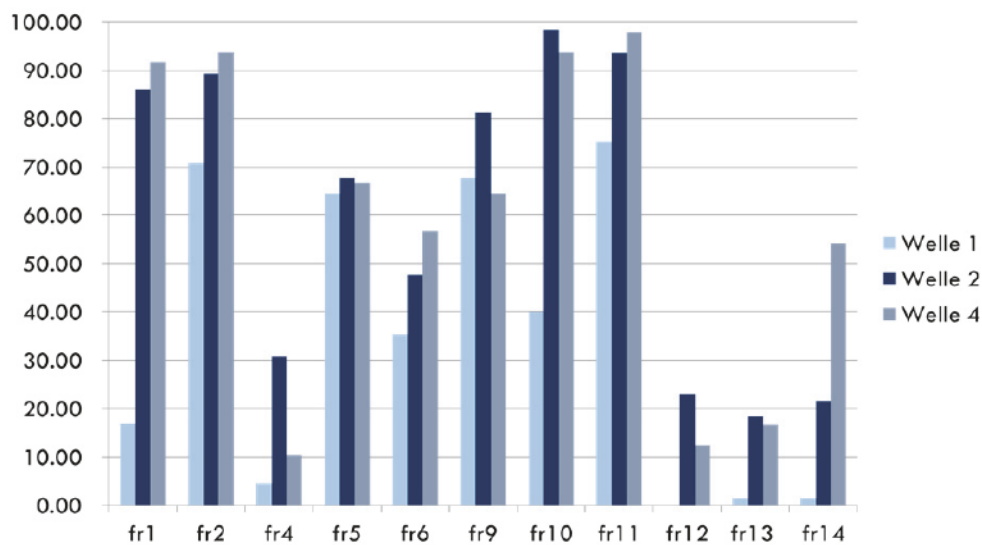


Abbildung 10: Wissensfragen zu TFD-Inhalten: Lösungshäufigkeit über 3 Messzeitpunkte

Unterrichten -Planen								
MZP 1 (Welle 1)			MZP 2 (Welle 2)			MZP 3 (Welle 4)		
Frage TFD1								
Häufigkeiten			Häufigkeiten			Häufigkeiten		
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	9	19,57%	richtig	43	93,48%	richtig	43	93,48%
falsch/ nicht gewusst	37	80,43%	falsch/ nicht gewusst	3	6,52%	falsch/ nicht gewusst	3	6,52%
Frage TFD2								
Häufigkeiten			Häufigkeiten			Häufigkeiten		
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	35	76,09%	richtig	43	93,48%	richtig	44	95,65%
falsch/ nicht gewusst	11	23,91%	falsch/ nicht gewusst	3	6,52%	falsch/ nicht gewusst	2	4,35%
Frage TEDS-M2.1								
Häufigkeiten			Häufigkeiten			Häufigkeiten		
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	43	93,48%	richtig	44	95,65%	richtig	46	100,00%
falsch/ nicht gewusst	3	6,52%	falsch/ nicht gewusst	2	4,35%	falsch/ nicht gewusst	0	0,00%
Frage TEDS-M2.2								
Häufigkeiten			Häufigkeiten			Häufigkeiten		
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	42	91,30%	richtig	44	95,65%	richtig	46	100,00%
falsch/ nicht gewusst	4	8,70%	falsch/ nicht gewusst	2	4,35%	falsch/ nicht gewusst	0	0,00%
Frage TEDS-M2.3								
Häufigkeiten			Häufigkeiten			Häufigkeiten		
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	44	95,65%	richtig	45	97,83%	richtig	46	100,00%
falsch/ nicht gewusst	2	4,35%	falsch/ nicht gewusst	1	2,17%	falsch/ nicht gewusst	0	0,00%
Frage TEDS-M2.4								
Häufigkeiten			Häufigkeiten			Häufigkeiten		
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	37	80,43%	richtig	39	84,78%	richtig	40	86,96%
falsch/ nicht gewusst	9	19,57%	falsch/ nicht gewusst	7	15,22,%	falsch/ nicht gewusst	6	13,04%



Frage TEDS-M2.5								
Häufigkeiten			Häufigkeiten			Häufigkeiten		
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	20	43,48%	richtig	17	36,96%	richtig	20	43,48%
falsch/ nicht gewusst	26	56,52%	falsch/ nicht gewusst	29	63,04%	falsch/ nicht gewusst	26	56,52,%
Frage TEDS-M4.1								
Häufigkeiten			Häufigkeiten			Häufigkeiten		
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	16	34,78%	richtig	44	95,65%	richtig	45	97,83%
falsch/ nicht gewusst	30	65,22%	falsch/ nicht gewusst	2	4,35%	falsch/ nicht gewusst	1	2,17%
Frage TEDS-M4.2								
Häufigkeiten			Häufigkeiten			Häufigkeiten		
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	17	36,96%	richtig	42	91,30%	richtig	42	91,30%
falsch/ nicht gewusst	29	63,04%	falsch/ nicht gewusst	4	8,70,%	falsch/ nicht gewusst	4	8,70,%
Frage TEDS-M4.3								
Häufigkeiten			Häufigkeiten			Häufigkeiten		
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	16	34,78%	richtig	43	93,48%	richtig	44	95,65%
falsch/ nicht gewusst	30	65,22%	falsch/ nicht gewusst	3	6,52%	falsch/ nicht gewusst	2	4,35,%
Frage TEDS-M4.4								
Häufigkeiten			Häufigkeiten			Häufigkeiten		
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	10	21,74%	richtig	27	58,70%	richtig	39	84,78%
falsch/ nicht gewusst	36	78,26%	falsch/ nicht gewusst	19	41,30%	falsch/ nicht gewusst	7	15,22,%

Tabelle 48: Ausführliche Testergebnisse im Bereich Unterrichten – Planen

Unterrichten - Umsetzen								
MZP 1 (Welle 1)			MZP 2 (Welle 2)			MZP 3 (Welle 4)		
Frage TFD12.1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	19	41,30%	richtig	33	71,74%	richtig	36	78,26%
falsch/ nicht gewusst	27	58,70%	falsch/ nicht gewusst	13	28,26%	falsch/ nicht gewusst	10	21,74%
Frage TFD12.2								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	16	34,78%	richtig	35	76,09%	richtig	28	60,87%
falsch/ nicht gewusst	30	65,22%	falsch/ nicht gewusst	11	23,91%	falsch/ nicht gewusst	18	39,13%
Frage TFD12.3								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	21	45,65%	richtig	39	84,78%	richtig	38	82,61%
falsch/ nicht gewusst	25	54,35%	falsch/ nicht gewusst	7	15,22%	falsch/ nicht gewusst	8	17,39%
Frage TFD12.4								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	13	28,26%	richtig	34	73,91%	richtig	31	67,39%
falsch/ nicht gewusst	33	71,74%	falsch/ nicht gewusst	12	26,09%	falsch/ nicht gewusst	15	32,61%
Frage TFD12.5								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	13	28,26%	richtig	27	58,70%	richtig	17	36,96%
falsch/ nicht gewusst	33	71,74%	falsch/ nicht gewusst	19	41,30%	falsch/ nicht gewusst	29	63,04%
Frage TFD14.1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	18	39,13%	richtig	29	63,04%	richtig	37	80,43%
falsch/ nicht gewusst	28	60,87%	falsch/ nicht gewusst	17	36,96%	falsch/ nicht gewusst	9	19,57%

Frage TFD14.2								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	13	28,26%	richtig	20	43,48%	richtig	36	78,26%
falsch/ nicht gewusst	33	71,74%	falsch/ nicht gewusst	26	56,52%	falsch/ nicht gewusst	10	21,74,%
Frage TFD14.3								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	12	26,09%	richtig	20	43,48%	richtig	30	65,22%
falsch/ nicht gewusst	34	73,91%	falsch/ nicht gewusst	26	56,52%	falsch/ nicht gewusst	16	34,78%
Frage TFD15								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	41	89,13%	richtig	44	95,65%	richtig	-	-
falsch/ nicht gewusst	5	10,87%	falsch/ nicht gewusst	2	4,35,%	falsch/ nicht gewusst	-	-
Frage TEDS-M1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	25	54,35%	richtig	37	80,43%	richtig	42	91,30%
falsch/ nicht gewusst	21	45,65%	falsch/ nicht gewusst	9	19,57%	falsch/ nicht gewusst	4	8,70,%
Frage TEDS-M6.1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	5	10,87%	richtig	10	21,74%	richtig	11	23,91%
falsch/ nicht gewusst	41	89,13%	falsch/ nicht gewusst	36	78,26%	falsch/ nicht gewusst	35	76,09%
Frage TEDS-M6.2								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	6	13,04%	richtig	7	15,22%	richtig	5	10,87%
falsch/ nicht gewusst	40	86,96%	falsch/ nicht gewusst	39	84,78%	falsch/ nicht gewusst	41	89,13%
Frage TEDS-M6.3								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ

richtig	3	6,52%	richtig	6	13,04%	richtig	3	6,52%
falsch/ nicht gewusst	43	93,48%	falsch/ nicht gewusst	40	86,96%	falsch/ nicht gewusst	43	93,48,%
Frage TEDS-M8								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	19	41,30%	richtig	34	73,91%	richtig	32	69,57%
falsch/ nicht gewusst	27	58,70%	falsch/ nicht gewusst	12	26,09%	falsch/ nicht gewusst	14	30,43%

Tabelle 49: Ausführliche Testergebnisse für den Bereich Unterrichten – Umsetzen



Leadership								
MZP 1 (Welle 1)			MZP 2 (Welle 2)			MZP 3 (Welle 4)		
Frage TFD3								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	31	67,93%	richtig	46	100,00%	richtig	-	-
falsch/ nicht gewusst	15	32,61%	falsch/ nicht gewusst	0	0,00%	falsch/ nicht gewusst	-	-
Frage TFD4.1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	5	10,87%	richtig	38	82,61%	richtig	33	71,74%
falsch/ nicht gewusst	41	89,13%	falsch/ nicht gewusst	8	17,39%	falsch/ nicht gewusst	13	28,26%
Frage TFD4.2								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	4	8,70%	richtig	36	78,26%	richtig	33	71,74%
falsch/ nicht gewusst	42	91,30%	falsch/ nicht gewusst	10	21,74%	falsch/ nicht gewusst	13	28,26%
Frage TFD4.3								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	3	6,52%	richtig	32	69,57%	richtig	14	30,43%
falsch/ nicht gewusst	43	93,48%	falsch/ nicht gewusst	14	30,43%	falsch/ nicht gewusst	32	69,57%
Frage TFD4.4								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	2	4,35%	richtig	17	36,96%	richtig	5	10,87%
falsch/ nicht gewusst	44	95,65%	falsch/ nicht gewusst	29	63,04%	falsch/ nicht gewusst	41	89,13%
Frage TFD5								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	28	60,87%	richtig	27	58,70%	richtig	32	69,57%
falsch/ nicht gewusst	18	39,13%	falsch/ nicht gewusst	19	41,30%	falsch/ nicht gewusst	14	30,43%

Frage TFD6								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	17	36,96%	richtig	27	58,70%	richtig	21	45,65%
falsch/ nicht gewusst	29	63,04%	falsch/ nicht gewusst	19	41,30%	falsch/ nicht gewusst	25	54,35%
MZIP1 (Welle 3)			MZIP2 (Welle 4)			MZIP3 (liegt nicht vor)		
Frage TFD16.1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	7	16,28%	richtig	11	25,58%	richtig	-	-
falsch/ nicht gewusst	36	83,72%	falsch/ nicht gewusst	32	74,42%	falsch/ nicht gewusst	-	-
Frage TFD16.2								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	7	16,28%	richtig	11	25,58%	richtig	-	-
falsch/ nicht gewusst	36	83,72%	falsch/ nicht gewusst	32	74,42%	falsch/ nicht gewusst	-	-
Frage TFD16.3								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	7	16,28%	richtig	11	25,58%	richtig	-	-
falsch/ nicht gewusst	36	83,72%	falsch/ nicht gewusst	32	74,42%	falsch/ nicht gewusst	-	-
Frage TFD16.4								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	7	16,28%	richtig	11	25,58%	richtig	-	-
falsch/ nicht gewusst	36	83,72%	falsch/ nicht gewusst	32	74,42%	falsch/ nicht gewusst	-	-
Frage TFD17								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	6	13,95%	richtig	6	13,95%	richtig	-	-
falsch/ nicht gewusst	37	86,05%	falsch/ nicht gewusst	37	86,05%	falsch/ nicht gewusst	-	-

Frage TFD18								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	24	55,81%	richtig	27	62,79%	richtig	-	-
falsch/ nicht gewusst	19	44,19%	falsch/ nicht gewusst	16	37,21%	falsch/ nicht gewusst	-	-
Frage TFD19								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	15	34,88%	richtig	16	37,21%	richtig	-	-
falsch/ nicht gewusst	28	65,12%	falsch/ nicht gewusst	27	62,79%	falsch/ nicht gewusst	-	-
Frage TFD20								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	22	51,16%	richtig	19	44,19%	richtig	-	-
falsch/ nicht gewusst	21	48,84%	falsch/ nicht gewusst	24	55,81%	falsch/ nicht gewusst	-	-

Tabelle 50: Ausführliche Ergebnisse im Bereich Leadership

Diagnose								
MZP 1 (Welle 1)			MZP 2 (Welle 2)			MZP 3 (Welle 4)		
Frage TFD10								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	20	43,48%	richtig	44	95,65%	richtig	44	95,65%
falsch/ nicht gewusst	26	56,52%	falsch/ nicht gewusst	2	4,35%	falsch/ nicht gewusst	2	4,35%
Frage TFD11								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	36	78,26%	richtig	41	89,13%	richtig	46	100,00%
falsch/ nicht gewusst	10	21,74%	falsch/ nicht gewusst	5	10,87%	falsch/ nicht gewusst	0	0,00%
Frage TEDS-M5.1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	31	67,39%	richtig	41	89,13%	richtig	42	91,30%
falsch/ nicht gewusst	15	32,61%	falsch/ nicht gewusst	5	10,87%	falsch/ nicht gewusst	4	8,70%
Frage TEDS-M5.2								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	41	89,13%	richtig	45	97,83%	richtig	43	93,48%
falsch/ nicht gewusst	5	10,87%	falsch/ nicht gewusst	1	2,17%	falsch/ nicht gewusst	3	6,52%
Frage TEDS-M7.1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	31	67,39%	richtig	32	69,57%	richtig	36	78,26%
falsch/ nicht gewusst	15	32,61%	falsch/ nicht gewusst	14	30,43%	falsch/ nicht gewusst	10	21,74%
Frage TEDS-M7.2								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	28	60,87%	richtig	35	76,09%	richtig	39	84,78%
falsch/ nicht gewusst	18	39,13%	falsch/ nicht gewusst	11	23,91%	falsch/ nicht gewusst	7	15,22%



Frage TEDS-M7.3								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	23	50,00%	richtig	29	63,04%	richtig	34	73,91%
falsch/ nicht gewusst	23	50,00%	falsch/ nicht gewusst	17	36,96%	falsch/ nicht gewusst	12	26,09%
Frage TEDS-M7.4								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	16	34,78%	richtig	16	34,78%	richtig	27	58,70%
falsch/ nicht gewusst	30	65,22%	falsch/ nicht gewusst	30	65,22%	falsch/ nicht gewusst	19	41,30%

Tabelle 51: Ausführliche Ergebnisse für den Bereich Diagnose

Sprachförderung								
MZP 1 (Welle 1)			MZP 2 (Welle 2)			MZP 3 (Welle 4)		
Frage TFD13.1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	1	2,17%	richtig	8	17,39%	richtig	11	23,91%
falsch/ nicht gewusst	45	97,83%	falsch/ nicht gewusst	38	82,61%	falsch/ nicht gewusst	35	76,09%
Frage TFD13.2								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	1	2,17%	richtig	12	26,09%	richtig	13	28,26%
falsch/ nicht gewusst	45	97,83%	falsch/ nicht gewusst	34	73,91%	falsch/ nicht gewusst	33	71,74%
Frage TFD13.3								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	1	2,17%	richtig	21	45,65%	richtig	19	41,30%
falsch/ nicht gewusst	45	97,83%	falsch/ nicht gewusst	25	54,35%	falsch/ nicht gewusst	27	58,70%
Frage TEDS-M3.1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	40	86,96%	richtig	41	89,13%	richtig	39	84,78%
falsch/ nicht gewusst	6	13,04%	falsch/ nicht gewusst	5	10,87%	falsch/ nicht gewusst	7	15,22%
Frage TEDS-M3.1.1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	36	78,26%	richtig	40	86,96%	richtig	27	58,70%
falsch/ nicht gewusst	10	21,74%	falsch/ nicht gewusst	6	13,04%	falsch/ nicht gewusst	19	41,30%
Frage TEDS-M3.2								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	15	35,71%	richtig	17	36,96%	richtig	22	47,83%
falsch/ nicht gewusst	27	64,29%	falsch/ nicht gewusst	29	63,04%	falsch/ nicht gewusst	24	52,17%

Frage TEDS-M3.2.1								
	Häufigkeiten			Häufigkeiten			Häufigkeiten	
Antwort	absolut	relativ	Antwort	absolut	relativ	Antwort	absolut	relativ
richtig	32	74,42%	richtig	28	63,64%	richtig	29	63,04%
falsch/ nicht gewusst	11	25,58%	falsch/ nicht gewusst	16	36,36%	falsch/ nicht gewusst	17	36,96%

Tabelle 52: Ausführliche Ergebnisse im Bereich Sprachförderung

## Materialien zur Bildungsforschung

Herausgegeben von der  
Gesellschaft zur Förderung Pädagogischer Forschung (GFPPF) und dem  
Deutschen Institut für Internationale Pädagogische Forschung (DIPF)

### **Band 1**

Renate Martini: „Schulautonomie“. Auswahlbibliographie 1989-1996. 1997. 220 S. ISBN 3-923638-17-5.)

### **Band 2**

Clive Hopes: Assessing, evaluating and assuring quality in schools in the European Union. 1998. 211 S. ISBN 3-923638-19-1.

### **Band 3**

Clive Hopes: Beurteilung, Evaluation und Sicherung der Qualität an Schulen in der Europäischen Union. 1998. 167 S. ISBN 3-923638-20-5.

### **Band 4**

Peter Döbrich, Ingrid Plath, Heinrich Trierscheid (Hrsg.): ArbeitsPlatz-Untersuchungen mit Hessischen Schulen. Zwischenergebnisse 1998. 1999. 272 S. ISBN 3-923638-21-3.

### **Band 5**

Hermann Avenarius / Hans Döbert (Hrsg.): „Schule in erweiterter Verantwortung“. Ein Berliner Modellversuch (1995 bis 1998). Abschlußbericht der wissenschaftlichen Begleitung. 1998. 89 S. ISBN 3-923638-22-1.

### **Band 6**

Peter Döbrich / Harry Neß (Hrsg.): EUROPASS-Berufsbildung – Anstoß und Projekt im nationalen Reformprozess –. Fachtagung am 2. Juni 1999. 2000. 156 S. ISBN 3-923638-24-8.

### **Band 7**

Peter Döbrich (Hrsg.): Qualitätsentwicklung im naturwissenschaftlichen Unterricht. Fachtagung am 15. Dezember 1999. 2002. 68 S. ISBN 3-923638-25-6.

### **Band 8**

Harry Neß / Peter Döbrich (Hrsg.): Doppeltqualifizierende Bildungswege – ein europäisches Modell für die Zukunft?! Fachtagung am 19. September 2001. 2003. 162 S. ISBN 3-923638-26-4.

### **Band 9**

Peter Döbrich / Bernd Frommelt (Hrsg.): Europäisierung und Reform der Lehrerbildung in Hessen und Rheinland-Pfalz. Jahrestagung am 26. und 27. März 2003. 2004. 78 S. ISBN 3-923638-27-2.

### **Band 10**

Brigitte Steinert / Marius Gerech / Eckhard Klieme / Peter Döbrich: Skalen zur Schulqualität: Dokumentation der Erhebungsinstrumente. ArbeitsPlatzUntersuchung (APU) / Pädagogische Entwicklungsbilanzen (PEB). 2003. 170 S. ISBN 3-923638-28-0.

### **Band 11**

Martina Diedrich / Hermann Josef Abs / Eckhard Klieme: Evaluation im BLK-Modellprogramm Demokratie lernen und leben: Skalen zur Befragung von Schüler/-innen, Lehrer/-innen und Schulleitungen. 2004. 189 S. ISBN 3-923638-29-9.

### **Band 12**

Hermann Josef Abs / Peter Döbrich / Erika Vögele / Eckhard Klieme: Skalen zur Qualität der Lehrerbildung – Dokumentation der Erhebungsinstrumente: Pädagogische Entwicklungsbilanzen an Studienseminaren (PEB-Sem). 2. überarb. Auflage. 2005. 125 S. ISBN 3-923638-30-2.

### **Band 13**

Katrin Rakoczy / Alex Buff / Frank Lipowsky: Teil 1: Befragungsinstrumente. In: Eckhard Klieme / Christine Pauli / Kurt Reusser (Hrsg.): Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“. 2005. 297 S. ISBN 3-923638-31-0.



**Band 14**

Frank Lipowsky / Barbara Drollinger-Vetter / Johannes Hartig / Eckhard Klieme: Teil 2: Leistungstests. In: Eckhard Klieme / Christine Pauli / Kurt Reusser (Hrsg.): Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“. 2006. 114 S. ISBN-10: 3-923638-32-9; ISBN 13: 978-3-923638-32-1.

**Band 15**

Isabelle Hugener / Christine Pauli / Kurt Reusser: Teil 3: Videoanalysen. In: Eckhard Klieme / Christine Pauli / Kurt Reusser (Hrsg.): Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“. 2006. 270 S. ISBN-10: 3-923638-33-7; ISBN-13: 978-3-923638-33-8.

**Band 16**

Marius Gerecht: Schulqualität und Schulevaluation – Schulspezifische Rückmeldung auf der Basis der Pädagogischen Entwicklungsbilanzen. 2006. 167 S. ISBN-10: 3-923638-34-5, ISBN-13: 978-3-923638-34-5.

**Band 17**

Marius Gerecht / Brigitte Steinert / Eckhard Klieme / Peter Döbrich: Skalen zur Schulqualität: Dokumentation der Erhebungsinstrumente. Pädagogische Entwicklungsbilanzen mit Schulen (PEB). 2. überarb. Auflage. 2007. 122 S. ISBN-10: 3-923638-35-3, ISBN-13: 978-3-923638-35-2.

**Band 18**

Peter Döbrich / Marius Gerecht / Jutta Laukart / Herbert Schnell: Skalen zur Qualität der Schulaufsicht: Dokumentation der Erhebungsinstrumente – Entwicklungsbilanzen im Schulamt (EBIS). 2007. 70 S. ISBN-10: 3-923638-36-1, ISBN-13: 978-3-923638-36-9.

**Band 19**

Hermann Josef Abs / Nina Roczen / Eckhard Klieme: Abschlussbericht zur Evaluation des BLK-Programms „Demokratie lernen und leben“. 2007. 86 S. ISBN: 978-3-923638-37-6.

**Band 20**

Hermann Josef Abs / Martina Diedrich / Helge Sickmann / Eckhard Klieme: Evaluation im BLK-Modellprogramm Demokratie lernen und leben: Skalen zur Befragung von Schüler/-innen, Lehrer/-innen und Schulleitungen. Dokumentation der Erhebungsinstrumente 2006. 2007. 124 S. ISBN: 978-3-923638-38-3.

**Band 21**

Peter Döbrich / Herbert Schnell (Hrsg.): QualitätsPartnerschaft der Regionen (QPR) – Europäische Indikatoren für Schulentwicklung und ihre Evaluation. 2008. 95 S. ISBN: 978-3-923638-39-0.

**Band 22**

Hermann Josef Abs / Peter Döbrich / Anne Gerlach-Jahn / Eckhard Klieme: Pädagogische Entwicklungsbilanzen an Studienseminaren (PEB-Sem). Auswahl und statistische Analyse der Erhebungsinstrumente. 2009. 154 S. ISBN 978-3-923638-40-6.

**Band 23/1**

Frank Lipowsky / Gabriele Faust / Karina Greb (Hrsg.): Dokumentation der Erhebungsinstrumente des Projekts „Persönlichkeits- und Lernentwicklung von Grundschulern“ (PERLE) – Teil 1. Karina Greb / Sebastian Poloczec / Frank Lipowsky / Gabriele Faust: PERLE-Instrumente: Schüler, Lehrer, Eltern (Messzeitpunkt 1). 2009. 184 S. ISBN 978-3-923638-41-3.

**Band 23/2**

Frank Lipowsky / Gabriele Faust / Karina Karst (Hrsg.): Dokumentation der Erhebungsinstrumente des Projekts „Persönlichkeits- und Lernentwicklung von Grundschulkindern“ (PERLE) – Teil 2. Karina Karst / Emely Mösko / Frank Lipowsky / Gabriele Faust: PERLE-Instrumente: Schüler, Eltern (Messzeitpunkte 2 & 3). 2011. 160 S. ISBN 978-3-923638-49-9.

**Band 24**

Holger Quellenberg: Studie zur Entwicklung von Ganztagschulen (StEG) — ausgewählte Hintergrundvariablen, Skalen und Indices der ersten Erhebungswelle. 2009. 154 S. ISBN 978-3-923638-42-0.

**Band 25/1**

Wolfgang Wagner / Andreas Helmke / Ernst Rösner: Deutsch Englisch Schülerleistungen International. Dokumentation der Erhebungsinstrumente für Schülerinnen und Schüler, Eltern und Lehrkräfte. 2009. 257 S. ISBN 978-3-923638-43-7.

**Band 26**

Bernd Frommelt / Marc Rittberger (Hrsg.): GFPP & DIPF. Dokumentation einer Kooperation seit 1950. Zusammenstellung und Redaktion: Peter Döbrich und Ulrich Schäfer. 2010. 272 S. ISBN 978-3-923638-44-4.

**Band 27**

Torsten Dietze: Zum Übergang auf weiterführende Schulen – Auswertung schulstatistischer Daten aus 10 Bundesländern. 2010. 129 S. ISBN 978-3-923638-46-8.

**Band 28**

Monika Buhl / Harm Kuper / Andrea Goldenbaum / Jana Höhler / Daniela Lindner / Stefan Müller-Mathis: Bericht zur Evaluation des Buddy-Landesprogramms in Hessen. 2011. 32 S. ISBN 978-3-923638-47-5.

**Band 29**

Klaus Hahne / Ulrich Schäfer: Das Projekt als Lehr-Lern-Form in der Berufsbildung in Deutschland. Eine Bibliographie für die Jahre 1956 bis 2010. 2011. 220 S. ISBN 978-3-923638-48-2.

**Band 30**

Ulrike Weyland / Eveline Wittmann: Expertise. Praxissemester im Rahmen der Lehrerbildung. 1. Phase an hessischen Hochschulen. 2011. 70 S. ISBN 978-3-923638-50-5.

**Band 31/1**

Peter Döbrich / Hartmut Storch: Pädagogische Entwicklungsbilanzen mit Studien-SEMinaren oder: Lehrerausbildung ohne Bilanzierung?. 2012. 162 S. ISBN 978-3-923638-51-2.

**Band 32**

Hermann Josef Abs / Ludwig Stecher / Julia Hohmann / Katrin Knoll / Katharina Golsch: Bericht zum Programmmonitoring 2012/2013 im Modellprogramm „Kulturagenten für kreative Schulen“. 2013. 87 S. ISBN 978-3-923638-53-6.

**Band 33**

Horst Weihaupt / Rosa Scherer / Mareike Tarazona / Marisa Richter / Rabea Krätschmer-Hahn / Karin Zimmer: Zur Situation kultureller Bildung an Schulen. Ergebnisse einer Schulleitungsbefragung 2011. 2013. 111 S. ISBN 978-3-923638-54-3.

**Band 34**

Hermann Josef Abs / Thomas Eckert / Eva Anderson-Park: Effektivität der Qualifizierung von Teach First Fellows. Abschlussbericht zur summativen Evaluation der Sommerakademie von Teach First Deutschland. 2016. 100 S. ISBN 978-3-923638-55-0.

Die Reihe wird fortgesetzt.

Alle Publikationen stehen Online als Volltext zur Verfügung unter: <http://www.pedocs.de>

